Radboud Repository

Radboud University Nijmegen

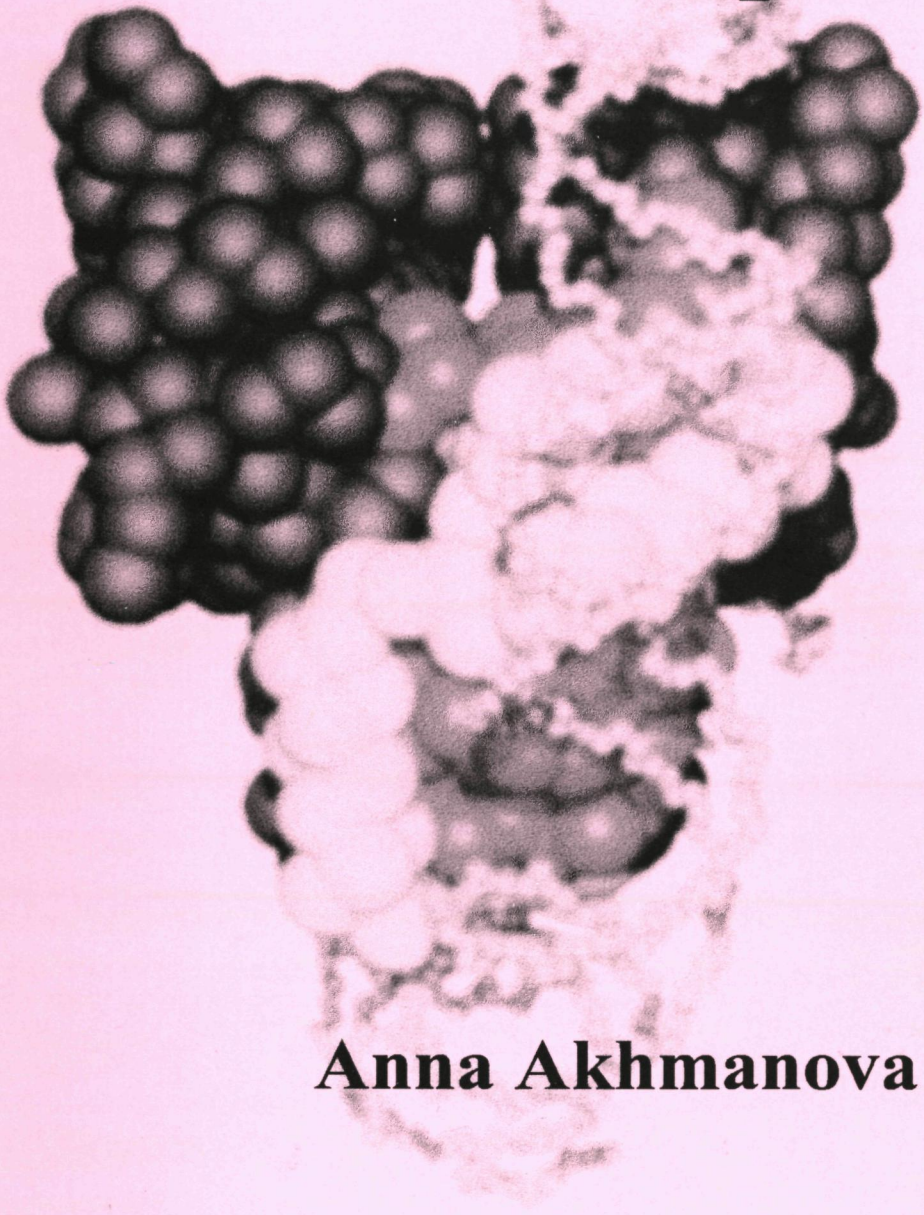
PDF hosted at the Radboud Repository of the Radboud University Nijmegen

The following full text is a publisher's version.

For additional information about this publication click this link.
http://hdl.handle.net/2066/146342

Please be advised that this information was generated on 2017-12-05 and may be subject to change.

Histone gene expression in *Drosophila*



Anna Akhmanova

HISTONE GENE EXPRESSION IN *DROSOPHILA*

een wetenschappelijke proeve

op het gebied van de Natuurwetenschappen

Proefschrift

ter verkrijging van de graad van doctor
aan de Katholieke Universiteit Nijmegen.
volgens besluit van het College van Decanen
in het openbaar te verdedigen
op dinsdag 4 maart 1997
des namiddags om 1.30 precies

door

Anna Sergeevna Akhmanova

geboren op 11 mei 1967
te Moskou

Promotor: Prof. Dr. W. Hennig

Manuscriptcommissie:

Prof. Dr. C. Heyting (Landbouwuniversiteit Wageningen)

Prof. Dr. H. Jäckle (Max Plank Institute of Biophysical Chemistry, Göttingen)

Prof. Dr. R.N.H. Konings (Katholieke Universiteit Nijmegen)

ISBN 90-9010261-2

Contents

Chapter I

Histone replacement variants: distinguishing features and possible functions

Anna Akhmanova and Wolfgang Hennig

The feeling on me grows and grows
That hardly anybody knows
If those are these or these are those.

A. A. Milne, *Winnie-the-Pooh,*
Lines written by a bear of very little brain.

Histones are highly conserved small basic proteins which constitute the elementary units of chromatin in the nuclei of all eukaryotic cells, the nucleosomes. In higher eukaryotes histones are encoded by multigene families, containing members of two types: replication dependent, or cell-cycle regulated histone genes and replication independent, or replacement histone genes. Most of the genes from both groups are expressed in many different tissues, but some tissue-specific genes of both types can be found. These are sometimes considered as a separate class of histone genes (Schümperli, 1986).

The distinction between the two types of histone genes is based on their structure, genomic organization, mode of regulation and the type of mRNA. It is generally accepted that the bulk of the histone proteins is synthesized in the S-phase of the cell-cycle and is incorporated into the chromatin during the DNA replication. Replication dependent, or cell-cycle regulated histone genes are responsible for this process. The expression of these genes in most eukaryotes is tightly coupled to the DNA synthesis (for review, see Schümperli, 1988; Osley, 1991). In contrast, the replacement histone genes display no strict regulation in relation to the cell-cycle (Wu and Bonner, 1981; Schümperli, 1986). Initial studies in vertebrates have shown that while replication dependent histone genes are multicopy, contain no introns and produce short non-polyadenylated mRNAs (Wells, 1986; Wells and McBride, 1989), replacement histone genes are single copy and resemble the rest of the protein-coding genes by the presence of introns and poly(A) tails in their mRNAs (Brush et al., 1985; Schümperli, 1986). Numerous recent studies of the structure and function of histone genes from different representatives of both animal and plant kingdoms have shown that histone genes of both types exist in most of the eukaryotic organisms. In this review we will focus on the structural and functional features of the replacement histone genes.

1. Regulation in the cell cycle

The main distinctive feature of the replacement histones is the relative independence of their expression on the phase of the cell cycle. The expression of the cell-cycle regulated histone genes is strongly induced in the beginning of the S-phase and is repressed upon the cessation of the DNA replication (for review, see Schümperli, 1988, Marzluff and Pandey, 1988, Osley, 1991). The expression of some replacement histone genes is also stimulated, at least at the level of mRNA, during the S-phase. This was demonstrated in cultured mammalian cells for the replacement variants $H1^{0}$ (Grunwald et al., 1991; Khochbin et al., 1991) and H3.3 (Hraba-Renevey and Kress, 1989). The developmental expression pattern of the *Drosophila* H2A.F/Z-like variant, H2AvD, follows the pattern of replication-dependent histone gene expression (van Daal and Elgin, 1992). Also in plants the highest expression of the replacement variants of histone H3 was observed in young shoots, similar to the replication-dependent histones (Kanazin et al., 1996).

In eggs and very early embryos of many species the production of histone

Table 1. Histone replacement variants: nomenclature and characteristic features.

Histone type	Replacement variant	Synonyms	Organism	Primary structure features	mRNA type	Introns
H1	$H1^{o}$		mammals	<50% identity with cell-cycle regulated H1 more basic, than H1	polyadenylated	absent
		H1D, H1E, $H1^{s}$, H5	*Xenopus*			
		H1b	trout			
	H5		birds (erythrocytes)	<50% identity with cell-cycle regulated H1 very basic	polyadenylated	absent
	H1M	B4, H1X	*Xenopus* (eggs, early embryos)	~30% identity with major H1 less basic, than H1	polyadenylated	present
H2A	H2AF.Z	H2A.F H2A.Z M1	vertebrates	~60% identity with cell-cycle regulated H2A	polyadenylated	present
		H2AvD	*Drosophila*			
		hv1	*Tetrahymena*			
	H2A.X		vertebrates	C-terminus different from cell-cycle regulated H2A	non-poly-adenylated and polyadenylated	absent
H3	H3.3		animals	97% identity with cell-cycle regulated H3	polyadenylated	present
	H3.III	H3.2	higher plants	97% identity with cell-cycle regulated H3	polyadenylated	present
	hv2		*Tetrahymena*	~88% identity with *Tetrahymena* major H3	polyadenylated	absent
H4	H4r		*Drosophila*	identical to *Drosophila* cell-cycle regulated H4	polyadenylated	present

proteins is not coupled to DNA replication. During oogenesis and maturation, large reservoirs of histone mRNAs (*Drosophila*, sea urchin) or proteins (amphibians) are accumulated to be used in early embryogenesis (Anderson and Lengyel, 1984; Kedes, 1979; Woodland, 1980). Some replacement histone variants are synthesized and accumulated in oocytes, as was shown for the *Xenopus* H1M (B4) (Dworkin-Rastl et al., 1994) and *Drosophila* H3.3 and H4r (Akhmanova et al., 1995, 1996).

The main distinction between the histone genes of the two types is the expression of the replacement histones albeit at a low level in non-S-phase cells. In Chinese hamster ovary cells such basal histone synthesis constitutes less than 10% of that during the S-phase (Wu and Bonner, 1981). All four core histones participate in the basal synthesis, although only for H2A and H3 distinctive basal or replacement sequence variants are found (Wu and Bonner, 1981). The pattern of the H2A variant synthesis in quiescent cells is different from that in the G1/G2 cells (Wu et al., 1982), but the synthesis of H1 and H3 variants is not (D'Incalci et al., 1986; Wu et al., 1982). H2A.X and H2A.Z variants are the only H2A histones, expressed in G1/G2 cells. However, in quiescent cells, similar to the S-phase, all H2A subtypes are produced and the level of the H2A.Z and H2A.X histones remains low (Wu et al., 1984). On the other hand, the replacement histone H3.3 variant is the only histone H3 subtype which is synthesized both in non S-phase and in non-dividing cells. The H3.3-encoding genes were shown to be induced when tumor cells switch to a terminal differentiation program (Krimer et al., 1993). Histone H3.3 accumulates in differentiated cells (Wu et al. 1983; Grove and Zweidler, 1984; Brown et al., 1985; Wunsch and Lough, 1987; Dell'Orco and Worthington, 1988). The same probably holds true for non-dividing somatic tissues of *Drosophila*, where the transcripts of replacement, but not the replication-dependent histone genes can be detected (Akhmanova et al., 1995, 1996).

Also in plants replacement histone H3 genes (H3.2 or H3.III) are expressed in cell-cycle independent manner and they produce the majority of the H3 mRNA and protein in non-meristematic tissues (Kapros et al., 1992, 1995; Kanazin et al. 1996). The steady-state level and the rate of turnover of the replacement histone H3 are generally much higher in plants than in animals, especially when growing cell populations are compared (Waterborg, 1992).

Establishment of the differentiated state in vertebrate cells correlates with the expression of the replacement type linker histone, $H1^{O}$ (for review, see Khochbin and Wolffe, 1994; Doenecke et al., 1994; Doenecke and Alonso, 1996). Like the H3.3, $H1^{O}$ is induced during in vitro terminal differentiation of tumor cell lines (reviewed by Zlatanova and Doenecke, 1994). A closely related tissue specific H1 variant, histone H5, accumulates in terminally differentiated avian erythrocytes (Sun et al., 1989).

Certain cell-cycle regulated histone genes also participate in basal histone synthesis. For example, the synthesis of mouse replication dependent histone variants H2A.1 and H2B.2 is completely repressed at the end of S-phase, while

replication dependent histones H2A.2 and H2B.1 retain a low expression level even when the DNA replication ceases (Zweidler, 1984).

The regulation of histone synthesis during the cell cycle occurs also in unicellular organisms, like, for example, in budding and in fission yeast (for review, see Osley, 1991). One can imagine that due to the absence of prolonged differentiated state, these organisms don't need a separate set of genes, insuring basal histone synthesis. That is probably the reason why replacement type histone genes are absent in budding yeast, *Saccharomyces cerevisiae*. Also in fission yeast, *Schizosaccharomyces pombe*, all histone genes are cell-cycle regulated, including the homologue of the animal H2A.F/Z gene, *pht1* (Carr et al., 1994).

A special case is represented by *Tetrahymena* and other ciliates. These protozoans contain two nuclei: mitotically dividing, transcriptionally almost inert micronuclei and amitotically dividing, transcriptionally active macronuclei. In *Tetrahymena* histones of each type are encoded by several genes (Bannon and Gorovsky, 1984; Nomoto et al., 1987; Thatcher et al., 1994; Shen et al., 1995; Liu et al., 1996). The pattern of expression of individual histone genes in the log phase growth conditions is different from that in non-growing (starving) cells, the latter representing basal histone synthesis (Bannon and Gorovsky, 1984). For example, hv2 is a histone H3 subtype, which is preferentially synthesized and deposited in the macronuclei of non-growing (starved) cells (Bannon et al., 1983). Apparently, it is the ciliate analogue of the animal H3.3 and plant H3.2 (H3.III) (Thatcher et al., 1994). *Tetrahymena* also contains an H2A.F/Z homologue, the hv1 variant, which is also preferentially accumulated in macronuclei (Bannon and Gorovsky, 1984; van Daal et al., 1990).

2. Transcriptional regulation

S-phase dependent transcriptional regulation of replication dependent histone genes was a subject of intensive research in a variety of organisms (for reviews, see Osley, 1991; Stein et al., 1996). The promoters of the histone genes of this type contain some elements, common for many RNA polymerase II-transcribed genes, such as the TATA-box, CCAAT-box, Sp1-box and the octamer sequence. In addition, a number of specific promoter elements, such as the H1-box (Coles and Wells, 1985) or the H4 element (Ramsey-Ewing et al, 1994), which contribute to the activation of a particular histone gene type, were identified. S-phase control elements were found not only in the upstream sequences, but also within the open reading frames of the histone genes (see Kaludov et al., 1996, and references therein).

Among the promoters of the replacement histone genes, those of the $H1^{O}$ and H5 genes were studied most extensively (for reviews, see Khochbin and Wolffe, 1994; Doenecke et al., 1994; Doenecke and Alonso, 1996). Interestingly, these two linker histone promoters share some elements with the cell-cycle regulated histones, such as an H4TF2 binding site, also present in histone H4 genes, and the H1 box (in the $H1^{O}$ promoter). This is not surprising since the $H1^{O}$ histone transcription is stimulated in the S-phase of the cell-cycle and the

H4 cell-cycle regulated element and the H1 box are involved in cell-cycle dependent regulation. Both $H1^0$ and H5 promoters contain the TATA-box and Sp1-binding sites, but in contrast to the promoters of the replication-dependent H1 genes they lack the CCAAT-box. Specific elements, like upstream conserved element (UCE) in the $H1^0$ promoter and GATA-1 binding sites in the H5 promoter are also present.

TATA-boxes, CCAAT-boxes and Sp1-boxes can usually be found also in the upstream regions of many but not in all sequenced genes encoding replacement core histone variants in vertebrates (Brush et al., 1985; Wells et al., 1987; Hatch and Bonner, 1990; Albig et al., 1995). Among these genes, only the promoter of the H2A.X-encoding gene was analysed in some detail. The transcription from this promoter is not dependent on the cell-cycle phase (Bonner et al., 1993). It contains a TATA-box and two CCAAT elements. The proximal of these two elements, together with some surrounding sequences, was found to be an important determinant of the H2A.X promoter activity. It binds nuclear factors, which also bind to the CCAAT box from the H2A.Z promoter, but not to the promoter of a cell-cycle regulated H2A gene (Ivanova et al., 1994a).

Some of the S-phase control elements, characteristic for the cell-cycle regulated genes, are clearly absent in their replacement counterparts. For example, the the α and Ω elements of the coding region activating sequence (CRAS), present in the mouse histone H3.2 gene, are mutated in the H3.3 genes (Kaludov et al., 1996). Remarkably, the α-element of CRAS coincides with the region of the H3 open reading frame, where most of the amino acid differences between the replacement H3.3 variant and replication-dependent H3.1/2 variants are localized (Bowman and Hurt, 1995). This is also one of the most variable regions in the H3 sequence, when the H3 histones from different kingdoms are compared (Thatcher et al., 1994).

Little functional data on the promoters of replacement histone genes from invertebrates is available. In Drosophila, promoters of H2AvD (van Daal et al., 1990), H3.3B (Akhmanova et al., 1995) and H4r (Akhmanova et al., 1996) genes contain no TATA-box, in contrast to the cell-cycle regulated histone genes (Matsuo and Yamazaki, 1989; Kremer and Hennig, 1990). The generic control sequences, such as H4 cell-cycle control element, were not found either. By sequence comparison, a putative promoter element was identified in H2AvD and H3.3B genes (Akhmanova et al., 1995). However it was not found in the H4r gene (Akhmanova et al., 1996).

3. mRNA polyadenylation status and regulation of mRNA stability

The animal cell-cycle regulated histone mRNAs are the only type of mRNAs, known so far, which terminate not with a poly(A) tail, but with a highly conserved stem-loop structure (Birnstiel et al., 1985; Marzluff, 1992). This structure is essential for the cell-cycle regulation of processing and stability (Schümperli, 1988, Marzluff and Pandey, 1988; Mowry and Steitz, 1988). The replacement histone genes, on the other hand, produce polyadenylated mRNAs, comparable to

the rest of the protein-coding genes (Schümperli, 1986; see Fig. 1). Processing through polyadenylation is not subject to cell-cycle regulation. In addition, stability of the polyadenylated messengers is not correlated with the DNA synthesis, like the stability of the mRNAs, ending with the stem-loop (Sittman et al., 1983). The polyadenylation status of the mRNA, therefore, is commonly used to distinguish the histone genes of the two types.

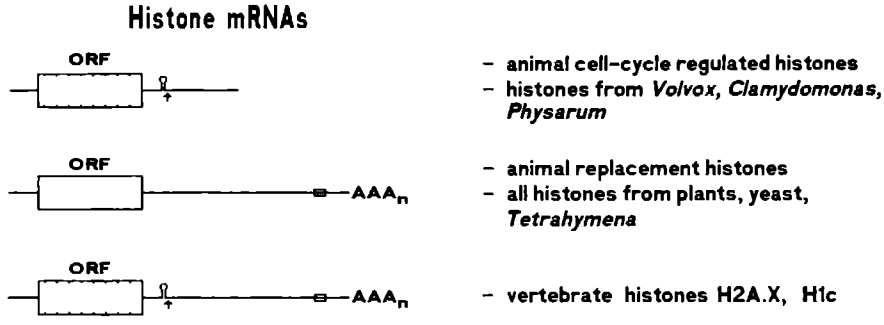


Fig. 1. Types of the 3' processing of histone mRNAs. Stippled boxes represent histone open reading frames, vertical arrows indicate cleavage sites after the stem-loop sequence; small open boxes represent polyadenylation signals.

This way of distinguishing between replication-dependent and replacement genes, though easy and convenient, has considerable limitations. Firstly, mRNAs of some histone genes can be processed via both pathways (see Fig. 1). This group of genes includes H2A.X of human and mouse (Mannironi et al., 1989; Nagata et al., 1991), which are regarded as replacement histones on the basis of their preferential synthesis during the G1 and G2 phases (Wu and Bonner, 1981; Wu et al., 1982). It also includes the chicken H1c-encoding genes H1.01 and H1.10 (Kirsh et al., 1989) and the mouse H1c (Cheng et al., 1989), which according to their amino acid sequence belong to the group of cell-cycle regulated linker histones.

Some non-polyadenylated histone mRNAs are synthesized in the absence of DNA synthesis. This is especially true for some testis-specific histones, like, for example, the mammalian H1t (Grimes et al., 1990; Drabent et al., 1991; for review, see Khochbin and Wolffe, 1994) or sperm-specific H2B subtypes in sea urchin (Busslinger and Barberis, 1985; Lai et al., 1986; Lai and Childs, 1986). These genes are expressed in the differentiating male germ cells after the DNA replication is completed. However, judging from the amino acid sequence, the genomic organization and the type of mRNA processing, these genes are more closely related to the cell-cycle regulated rather than to the replacement histones. Testis-specific histone genes with polyadenylated mRNAs exist as well, as for example the mouse spermatid-specific histone H2B (Moss et al., 1989). In addition, normal cell-cycle regulated histone genes can be expressed in vertebrate postmeiotic male germ cells (in spermatids) as both non-polyadenylated and polyadenylated mRNAs (Challoner et al., 1989; Moss et al., 1994).

Finally, all histone mRNAs of many lower eukaryotes, including budding and fission yeast and ciliates, as well as those of higher plants, are polyadenylated. In these organisms the histone mRNA processing is not controlled by the cell cycle, while the stability of histone mRNAs is apparently cell-cycle regulated. For example, in yeast, cell-cycle dependent periodic fluctuations of the histone mRNA levels were observed even when the contribution of transcriptional control was eliminated (reviewed by Osley, 1991). In higher plants the mRNAs of cell-cycle regulated histone H3.1 genes are less stable and more sensitive to the DNA replication inhibitors than those of the replacement-type H3 genes (Kapros et al., 1995). At present, the determinants of histone mRNA stability in these organisms are poorly understood.

Another interesting feature distinguishing the animal replacement histone mRNAs from the cell-cycle regulated ones is the length of the 3' untranslated regions (UTRs). The 3' UTRs of the non-polyadenylated cell-cycle dependent mRNAs are usually very short (less than 100 nt) and with the exception of the stem-loop sequence, are not conserved in evolution (Wells, 1986; Wells and McBride, 1989). In the replacement histone genes the 3' UTRs are usually much longer (hundreds of nucleotides) and are in some cases (like the vertebrate histone H3.3 mRNAs) extremely conserved over long evolutionary periods (Wells et al., 1987; Chalmers and Wells, 1990). The functions of these particular sequences are unknown. It is, however, becoming increasingly clear, that the eukaryotic 3' UTR sequences control mRNA translation and stability (Jackson and Standart, 1990; Jackson, 1993).

4. Gene structure

Polyadenylation status of histone mRNAs in animals usually correlates with the presence or absense of introns in the respective genes. It is known that at least in mammalian cells splicing and polyadenylation are coupled to a certain extent (see Nesic and Maquat, 1994, and references therein). Therefore it seems quite logical that replacement histone genes, that are expressed as polyadenylated messengers, contain introns. This holds true for all animal histones H3.3 (Brush et al., 1985; Wells et al., 1987; Akhmanova et al., 1995; Albig et al., 1995), histone H2A.F/Z (van Daal et al., 1990) and maternal linker histone H1M (B4) from *Xenopus* (Cho and Wolffe, 1994). However, this rule is not very strict, since polyadenylated mRNAs of $H1^{0}$ and H5 are transcribed from genes without introns (Krieg et al., 1983; Doenecke and Tönjes, 1986).

The cell-cycle regulated histone genes usually lack introns. This is true for the animal kingdom as well as for yeasts and higher plants. In mammals it was shown that the presence of an intron suppresses the stem-loop dependent 3' end processing and stimulates the utilization of cryptic downstream polyadenylation signals (Pandey et al., 1990). It is not surprising, therefore, that non-polyadenylated cell-cycle regulated histone mRNAs in animals are produced exclusively from intronless genes (Wells, 1986; for the most recent compilation of histone sequences, see Baxevanis and Landsman, 1996). This rule cannot be gene-

ralized for all eukaryotes. The histone genes of *Physarum* and *Volvox* as well as some histone genes of *Chlamydomonas*, contain introns, even though they are cell-cycle regulated and generate non-polyadenylated mRNAs ending with a stem-loop, similar to that of animal histones (Wilhelm and Wilhelm, 1989; Müller and Schmitt, 1988; Müller et al., 1990; Fabry et al., 1995)

The correlation between the replication independence of expression and the presence of introns apparently holds true for higher plants. Extensive studies of the family of the histone H3-encoding genes have shown that all the intron-containing members encode a replacement type histone H3 (H3.2 or H3.III) as can be judged from the protein sequence and the mode of expression (Chaubet et al., 1992, Kanazin et al., 1996; Robertson et al., 1996).

5. Gene organization

In higher eukaryotes cell-cycle regulated histone genes are usually present in multiple copies, which are to a certain extent clustered. Even in budding yeast, where only two copies of each core histone gene are present, they are arranged in two H2A-H2B and two H3-H4 pairs (for review, see Grunstein et al, 1984). In some species all replication-dependent histone genes are present in large clusters of relatively homogeneous repeats, where each repeat contains genes for all core histones and the histone H1. Such an organization is observed, for example, in *Drosophila* (Lifton et al., 1978; Matsuo and Yamazaki, 1989), *Notophthalmus* (Stephenson et al., 1981) and for the early histone genes of sea urchins (Kedes, 1979). In other species randomly assorted clusters of replication dependent histone genes are found. This is true for birds and mammals (for recent reports, see Tönjes et al., 1989; Gruber et al., 1990; Albig et al., 1993; Drabent et al., 1993; Dong et al., 1994; Brown, V.D. et al., 1996) as well as for the late genes of sea urchin (Kedes, 1979). It was shown recently that at least in mammals, such random clusters have a tendency to be associated, so that, for example, all human cell-cycle regulated H1 genes are present on the same chromosome (Albig et al., 1993).

The distinguishing feature of the replacement histone genes is that they are present in one or two copies and not linked to the cell-cycle regulated genes or to each other. This is well illustrated by the distribution of histone genes in *Drosophila*, where all cell-cycle regulated genes are clustered together near the centromere of the chromosome 2 (Lifton et al., 1978), while the two H3.3-encoding genes are localized on the X chromosome and the left arm of the chromosome 2 (Akhmanova et a., 1995) and the histone H2AvD and H4r replacement genes are present at two different locations on the third chromosome (van Daal and Elgin, 1992; Akhmanova et al., 1996). A quite similar situation is observed in mammals. The human histone $H1^{O}$, H2A.X, H2A.Z and H3.3-encoding genes are not linked to the cell-cycle regulated histone genes and are all localized on different chromosomes (Albig et al., 1993; Ivanova et al., 1994b; Popescu et al., 1994; Albig et al., 1995). The isolated genomic position of replacement histone genes in these organisms can be regarded as a distinctive feature of this type

of genes.

The situation is somewhat different in higher plants, where the distinction in gene organization between cell-cycle regulated and replacement histones is less pronounced. Firstly, the clustering of cell-cycle regulated histones genes is observed only occasionally, and these genes are distributed over different chromosomes (Kanazin et al., 1996; for review, see Chaboute et al., 1993). On the other hand, replacement genes, at least those encoding histone H3 variant, can be present in several copies, which are in some cases (in *Arabidopsis*) tightly linked with each other (Kanazin et al., 1996; Robertson et al., 1996; Chaubet et al., 1992). The general tendency that the copy number of cell-cycle regulated genes is considerably higher than that of replacement genes holds true also for plants (Robertson et al., 1996; Chaboute et al., 1993). This apparently reflects the requirement, general for all higher organisms, to ensure very high rate of histone synthesis during the S-phase.

6. Amino acid sequence differences between replication-dependent and replacement histones

The proteins, encoded by the replacement histone genes, were initially identified as histone variants on acetic acid-urea-Triton gels due to their mobility, which differed from that of the major histone subtypes (Panyim and Chalkey, 1969; Franklin and Zweidler, 1977; West and Bonner, 1980). In most cases the replacement histone proteins are different from their cell-cycle regulated counterparts in at least a few amino acid positions. The amount of such differences usually correlates with the degree of conservation of the particular histone protein in evolution. The linker histone, H1, is the least conserved of all histones (Cole, 1987; Schulze and Schulze, 1995). For example, the structures of the vertebrate histone $H1^{O}$ and H5 variants are significanly different from that of the replication dependent histone H1 (Khochbin and Wolffe, 1994; Doenecke and Alonso, 1996). These histones are shorter (190 versus 210 amino acids), and are quite divergent in the sequence of their N- and C-terminal tails. Also in the central globular region, which is conserved between all linker histones, the degree of identity is not high (Schulze and Schulze, 1995). The distinguishing feature of these differentiation-specific linker histones is their highly basic character, compared to the replication-dependent histones H1. This is especially apparent for the histone H5, where arginine replaces a high proportion of the lysine residues, present in the C-terminal domain of other H1 histones (Doenecke and Tönjes 1986). The cleavage linker histone, H1M(B4), found in *Xenopus*, is also considerably different from the somatic linker histones (29% of identity with *Xenopus* H1, Smith et al., 1988) and it is less basic than other histone H1 subtypes (Dimitrov et al., 1993).

Histone H2A displays the highest degree of sequence divergence among the core histones. The H2A.F/Z variant, found in a large variety of species, is only ~60% identical to the cell-cycle regulated H2A protein in a given species, with the differences distributed throughout the whole sequence (Hatch and Bonner,

1988). In addition, the *Drosophila* H2A.F/Z-type histone, H2AvD, is 16 amino acids longer in the C-terminus than its cell-cycle regulated counterpart (van Daal et al., 1988), while the *Tetrahymena* H2A variant, hv1, has a longer amino terminus (Allis et al., 1986). Vertebrates have an additional H2A replacement variant, H2A.X. The core region of this histone is nearly identical to the major vertebrate H2A, such as H2A.1. Its C-terminal region is longer than that of H2A.1 and has some similarity to the C-termini of the H2A proteins of lower eukaryotes (Mannironi et al., 1989).

Histone H3 is highly conserved in evolution. In agreement with this observation, its replacement variants differ from their cell-cycle regulated counterparts only by a few amino acid substitutions (see alignment in Thatcher et al., 1994). Only four substitutions distinguish the replacement and cell-cycle regulated histones H3 in animals (positions 31, 87, 89 and 90) and in plants (positions 31, 41, 87 and 90). While the positions of three of these four substitutions are the same in plants and animals, the amino acids themselves are conserved between animals and plants only in the cell-cycle regulated, but not the replacement histones H3. The replication-dependent and replacement H3 variants in *Tetrahymena* are more divergent (16 substitutions). Interestingly, three of these substitutions are at the same positions 31, 87 and 89, as in the animal H3 variants.

Finally, for the evolutionary most conserved histone, H4, a replacement gene was found only in *Drosophila* (Akhmanova et al., 1996) and in *Ascaris* (Duda et al., unpublished; EMBL accession numbers Z69289, Z69290). In *Drosophila*, this gene encodes a protein which is identical to the cell-cycle regulated H4 in the same species. This observation accentuates the point that replacement histone genes do not necessarily encode replacement *variant* histones, distinguishable by their amino acid sequence. This conclusion provides a basis for searching for replacement histone H2B genes, which so far were not found in any organism.

7. Gene evolution

While the genes for all five cell-cycle regulated histones are isolated in many different organisms, in no case a full complement of replacement histone genes was described. The replacement genes for histones H1, H2A and H3 were described for a large variety of organisms, but histone H4 replacement genes were found so far only in two classes of invertebrates (see above) and histone H2B replacement genes were not described at all. However, in spite of the incomplete data, some general conclusions about the origin and evolution of replacement histones can be made.

Two main criteria which are used in such an analysis are the protein sequences of the replacement histones and, in cases when they are present, the positions of the introns in the open reading frame. Using both of these criteria it was concluded that histone H2A replacement variant, H2A.F/Z, has arisen very early in eukaryotic evolution, as it is found in vertebrates, invertebrates and *Tetrahymena* (van Daal et al., 1990; see Fig. 2). An intronless gene, *pht1*, from *S. pombe* apparently also belongs to this group of the H2A variants judging from

its protein sequence, although strictly speaking it cannot be regarded as a replacement gene, because its expression is cell-cycle dependent (Carr et al 1994)

Histone H3 replacement variants are described in *Tetrahymena* animals and plants Comparison of the amino acid sequences and the localization of their introns indicates that these variants have arisen independently in these three groups (Thatcher et al , 1994, see Fig 2) On the other hand within each kingdom replacement histones H3 are highly conserved For example, the animal replacement histone H3 3 is identical in all vertebrate and invertebrate species analysed so far and differs from the animal cell-cycle regulated H3 variants Positions of some of the introns in its open reading frame are conserved between vertebrates and invertebrates (Akhmanova et al 1995) The same is true for the H3 III (H3 2) variants of higher plants (Kanazin et al , 1996 Robertson et al , 1996) Apparently the separation of replacement and cell-cycle regulated histones occurred early in both plant and animal evolution

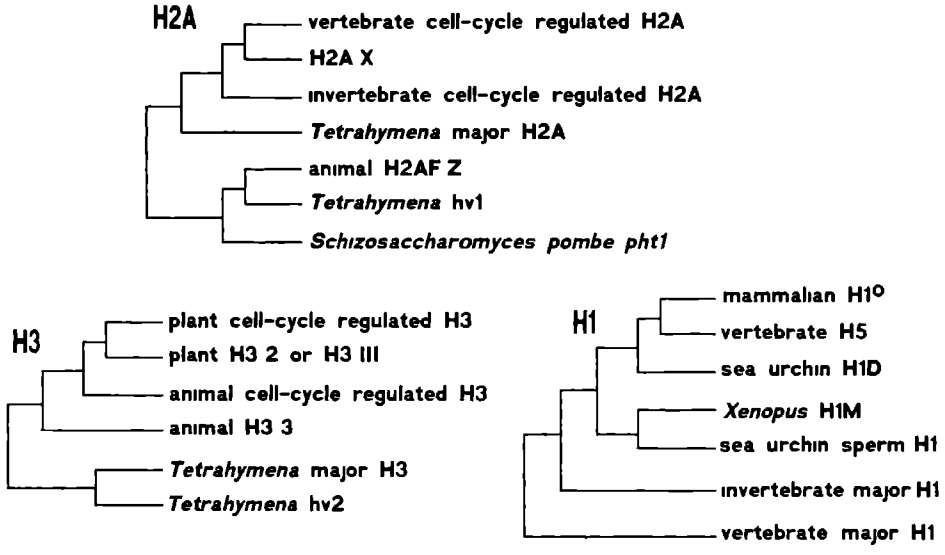


Fig 2 Phylogenetic trees calculated from the alignments of the amino acid sequences of histones H2A, H3 and globular domains of linker histones by neighbor-joining method The drawings represent the topologies of the phylogenetic trees illustrating the relationship between the replacement variants and cell-cycle regulated histones The branches are not drawn to scale and each branch represents in most cases a group of amino acid sequences These trees are simplified representations of phylogenetic trees for histones H2A and H3 calculated by Thatcher et al (1994) and for linker histones calculated by Schulze and Schulze (1995)

Similar to H3 3 variant, the replacement histones $H1^{0}$ from vertebrates (and related to it avian histone H5) are most closely related to replacement type his-

tone H1 from invertebrates (exemplified by the H1D histone from sea urchin, which is encoded by a polyadenylated mRNA; Schulze and Schulze, 1995). A phylogenetic tree deduced from the comparison of the conserved central domains of all linker histones places $H1^{0}$ and H5 together with H1M in close vicinity to invertebrate sperm histone H1 proteins and invertebrate histones H1, encoded by polyadenylated mRNAs (Schulze and Schulze, 1995; see Fig. 2). All these linker histones are more similar to invertebrate and plant replication-dependent histones H1 than to the vertebrate ones. Taken together, the data suggest an early evolutionary origin of the animal replication-independent linker histones.

Some replacement variants are of a more recent origin, as is clearly the case for the histone H2A.X. This histone is only found in vertebrates. It is more closely related to the replication-dependent histones H2A from vertebrates, than to the replacement variant H2A.F/Z or to invertebrate H2A histones (Mannironi et al, 1989; Thatcher and Gorovsky, 1994). As was mentioned before, the H2A.X-encoding genes contain processing signals characteristic for replication-dependent histones (the stem-loop) along with polyadenylation signals, characteristic of replacement histone genes. Therefore, this variant is likely to have originated from a cell-cycle regulated histone H2A gene in the course of the vertebrate evolution (see Fig. 2).

Taken together, the data described above show that replacement histones represent an evolutionary heterogeneous group of genes. An important feature, which might influence their evolution in higher eukaryotes as compared to the cell-cycle regulated histones is their solitary genomic position. This might contribute to the exclusion of these genes from the process of homogenization with other members of the multigene family. The evolutionary constraints on the replacement histone genes are apparently quite high, because replacement histones of a particular class (like H2AF/Z or H3.3) are often more conserved between different species, than their cell-cycle regulated counterparts (van Daal et al., 1990; Thatcher et al., 1994).

8. Functional differences between replication dependent and replacement histones

The central question in the study of the replacement histone genes concerns their functional significance. The presence of the histone genes of this type in all higher eukaryotes together with the high degree of their conservation in the course of evolution (see above) indicate that these genes are functionally important. Two possible functions, not necessarily mutually exclusive, can be proposed. On the one hand, these genes may insure chromatin maintenance over the long periods when cells do not divide but remain transcriptionally active (like most differentiated somatic cells in multicellular organisms). Nucleosomes are likely to become displaced and lost or damaged during transcription, and the repair process would require some histone synthesis. Such kind of synthesis can be best described as "basal" and does not have to involve histone proteins, which are different in their structure from the cell-cycle regulated histones. It is known that such basal synthesis can be also provided by at least some cell-

cycle regulated genes (see above). However, if the level of replication-independent expression of cell-cycle regulated genes is low, the presence of histone genes regulated in a cell-cycle independent manner might be essential. The *Drosophila* histone H4 replacement gene, H4r, fits well into such a concept, since it encodes a protein identical to the *Drosophila* cell-cycle regulated H4.

On the other hand, the replacement histones may have a truly "replacing" function by providing a source of histone proteins that can modify the nucleosome structure. The existence of some extreme histone variants, such as a histone H3-like centromeric protein CENP-A (Sullivan et al., 1994; Stoler et al., 1995) or the macro H2A histone, an H2A subtype containing a large C-terminal non-histone domain (Pehrson and Fried, 1992), indicates that at least some histone variants have special functions (for review, see Wolffe and Pruss, 1996b).

Such specialized functions are easy to imagine for some tissue-specific variants. For example, linker histones in general are thought to be involved in formation of higher order chromatin structure and in transcriptional repression (reviewed by Zlatanova and van Holde, 1996; Paranjape et al., 1994). Histone H5 is associated with transcription silencing in avian erythrocytes (Sun et al., 1989). It is more basic than other linker histones and might be more efficient in stabilization of the chromatin sructure. This is also true for the sea urchin sperm specific histones H1 and H2B with, which are involved in the transcriptional repression and genome packaging in sperm (for review, see Poccia, 1986; Romano, 1992). These histones contain highly basic tetrapeptide repeats, which are subject to reversible phosphorylation. In mature sperm these histones are present in a dephosphorylated state, which correlates with a high degree of chromatin condensation.

Cleavage linker histone from *Xenopus*, H1M (B4) illustrates an opposite tendency. It is less basic than other histone H1 subtypes (see above) and it is expected to display a weaker interaction with chromatin (Khochbin and Wolffe, 1994). This might facilitate frequent disruption of chromatin structure, which occurs during short cell cycles, characteristic for cleavage.

The question about functional significance becomes more difficult, however, when one considers ubiquitous replacement variants, such as H2A.F/Z or H3.3. Especially in the case of the H3 histone, where only a few amino acid substitutions distinguish replacement variants from the cell-cycle regulated histones, it is difficult to say if the observed differences are of functional importance or simply reflect neutral polymorphisms. The second interpretation, though attractive because of the small number of the observed changes, must however be treated with caution. It is becoming increasingly clear that core histones interact directly with a large number of chromatin proteins and that nucleosomes play an important role in gene regulation. Even small non-lethal changes in the histone structure can have very specific effects on the regulation of particular genes (Grunstein et al., 1992; Kruger et al., 1995; Wolffe and Pruss, 1996a).

Several approaches can be used to analyse the functional differences between histones variants. These approaches will be summarized below.

8.1. Influence on nucleosome and chromatin structure

8.1.1. Core histones

At present few direct data are available on the properties of nucleosomes and nucleosome arrays containing different histone subtypes. Early studies have shown that nucleosome core particles prepared from sea urchin sperm are more stable to heat denaturation and DNase digestion than those from chicken erythrocyte nuclei (Simpson and Bergman, 1980). This correlates nicely with the highly basic character of sea urchin sperm-specific histones. When nucleosome cores, containing early and late sea urchin histones (prepared from blastulae and plutei, respectively), were compared by similar experiments, the latter were found to be more stable (Simpson, 1981). This indicates that the type of core histones may have an impact on the nucleosome structure, although the influence of covalent modifications in this case cannot be ruled out.

Unfortunately, in most cases nucleosomes, prepared from natural sources, contain a mixture of histone variants. The situation is further complicated by multiple covalent modifications, some of which might have considerable impact on the nucleosome structure. One of the ways to overcome these problems is to analyse nucleosomes and chromatin fibers assembled in vitro from pure recombinant histones.

The influence of different histone isoforms on the nucleosome stability was also studied by hydroxyapatite dissociation chromatography (Li et al., 1993). This approach is based on the idea that histones, which destabilize the nucleosomes, will elute from the hydroxyapatite-bound chromatin at a lower salt concentration, than histones which are tightly bound. By this method it was shown that replacement histone variant H2A.Z as well as a replication-dependent H2B.2 stabilize the nucleosomes (Li et al., 1993)

8.1.2. Linker histones

Linker histones are involved in the organisation of chromatin in compact structures such as a 30-nm chromatin fiber (Thoma et al., 1979; for a recent review, see Zlatanova and van Holde, 1996). Several in vitro studies have addressed the differences between linker histone subtypes in binding to naked DNA or to chromatin. For example, it was shown that histone H5 and sperm-specific histone spH1 from sea urchins display cooperative binding to DNA at a salt concentration, lower than that found for a typical somatic H1 (Clark and Thomas, 1988). Average repeat length is different in chromatin, containing different histone H1 variants, which is presumably due to the variations in the length of the linker DNA. A correlation was observed between the histone spacing in H1-, H5- and spH1-DNA complexes and the linker length in H1-, H5- and spH1-containing chromatin (Clark and Thomas, 1988). However, replacement of histone H1 by histone H5 in cultured cells does not alter the nucleosomal repeat length in chromatin (Sun et al., 1990). This indicates that the linker histone is not solely responsible for nucleosome spacing.

To compare different histone H1 variants in their ability to condense chro-

matin in vitro, H1-depleted chromatin fragments were reassociated with purified individual linker histone subtypes (Biard-Roche et al., 1982; Marion et al., 1985). Reconstitution of chromatin with pure $H1^0$ gave rise to a looser, more nuclease-sensitive structure, than reconstitution with the major histone H1. In agreement with these data, histone H5, which is closely related to the $H1^0$ variant, was less effective in forming higher order chromatin structures than histone H1, as can be judged by aggregation/precipitation of linker histone-depleted chromatin (Nagaraja et al., 1995). However, in the same study it was shown that the trout equivalent of $H1^0$, H1b, is more effective in condensing chromatin, than the trout major histone H1 variant, H1a.

Studies of native chromatin have shown that the nuclease resistant, presumably more compact regions have elevated amounts of histone H1b in trout (Davie and Delcuve, 1991) or $H1^0$ in mammals (Roche et al., 1985a). The latter observation contradicts in vitro data on the decreased capacity of the histone $H1^0$ for chromatin condensation. This was interpreted as a result of the interaction of $H1^0$ with other chromosomal proteins. An alternative explanation is that histone $H1^0$, as well as histone H5, has greater affinity for preformed binding sites in higher order chromatin structures, than to the sites in H1-depleted chromatin. Competition experiments have shown that histone H5 preferentially binds to the native chromatin fragments while histone H1 does not distinguish between binding sites in native and H1-depleted chromatin (Thomas and Rees, 1983). It is possible, therefore, that histones $H1^0$ and H5 substitute histone H1 already bound in chromosome fibers with the consequence of a higher degree of condensation.

8.2. Expression level and accumulation in chromatin.

The expression level of different replacement histones is highly variable. Animal histone H3.3 is continuously expressed in quiescent cells. It gradually replaces the cell-cycle regulated histones, so that it can become a predominant H3 subtype in certain tissues (Zweidler, 1984). For example, in rat brain cortical neurons, the proliferation of which is restricted to a short period in brain development, histone H3.3 represents about 20% of total histone H3 at birth, but increases to nearly 90% at day 400 (Pina and Suau, 1987). This indicates that cells can tolerate gross changes in the content of this protein without significant alteration of their differentiation and metabolic status. One cannot exclude, of course, that H3.3 is a sort of an "ageing marker" or that its incorporation has an impact only in particular nucleosomes (like the nucleosomes bound to promoters).

Similar to H3.3, histone $H1^0$ has a tendency to accumulate in differentiated tissues (for review, see Khochbin and Wolffe, 1994; Zlatanova and Doenecke, 1994; Doenecke and Alonso, 1996). During prenatal mouse development $H1^0$ has been detected only in postmitotic lens fiber cells and yolk-sac-derived erythrocytes (Gjerset et al., 1982). After birth, $H1^0$ gradually accumulates in many tissues, reaching 25-30% of the total linker histone (Lennox and Cohen, 1983). It

was suggested that histone $H1^0$, due its highly basic character, might contribute to stabilization of particular chromatin structures, accompanying the arrest of cell proliferation (Khochbin and Wolffe, 1994). However, cells with significant amounts of $H1^0$ can still initiate DNA replication (Gorka et al., 1995). Therefore, low concentration of $H1^0$ is not a necessary prerequisite of cell proliferation.

The level of the histone H2A.F/Z remains relatively constant (~5-10% of the total H2A) and it never becomes the major H2A in any cell type (West and Bonner, 1980; Wu and Bonner, 1981; Wu et al., 1982). This is consistent with the localization of this variant in particular chromatin domains (see below). The assumption of special properties of the H2A.Z-containing nucleosomes is supported by the observation that the H2A.Z level was reduced in late stage trout testis chromatin, compared to the total H2A level (Nickel et al., 1987). This suggests that the H2A.Z-containing nucleosomes are selectively disassembled at certain stages of spermatogenesis in trout.

Histone H2A.X also represents about 10% of total H2A in most mammalian tissues and cell lines (Wu and Bonner, 1981; Wu et al., 1982). This variant increases considerably during mouse spermatogenesis (Bhatnagar et al., 1985). Apparently in this case its expression is related to a particular differentiation stage. High levels of the H2A.X histone were also observed in the *Xenopus* eggs and in the sperm chromatin of *Xenopus* during the first stages of decondensation in egg extracts (Dimitrov et al., 1994).

In plants, replacement histone H3 variant protein (H3.2 or H3.III) is highly expressed. The steady-state level of this variant protein displays a correlation with the size of the plant genome. H3.2 (H3.III) is a minor variant in tobacco, but it is more prominent in plants with smaller genomes (~30% in alfalfa) and it is the major histone H3 variant in *Arabidopsis* (Waterborg, 1992). This correlation was interpreted as an indication for the accumulation of this variant in transcriptionally active chromatin, since in plants with smaller genomes a larger proportion of the genome is transcriptionally active. Analysis of the turnover of the histone H3 variants in alfalfa tissue culture cells has shown that a considerable portion (more than 50%) of the replacement histone H3 is highly labile (Waterborg, 1993). High turnover of this histone H3 variant probably reflects the increased need for assembly of replacement nucleosomes in plant chromatin, compared to animal systems. Surprisingly, although histones H3 and H4 are incorporated into chromatin together as a heterotetramer, no corresponding high-turnover fraction is observed in the pool of the histone H4 in the same cells (Waterborg, 1993).

8.3. Distribution in the chromatin

The distribution of replacement variants between transcriptionally active and inactive chromatin was analysed in different systems. A low-salt-soluble polynucleosome fraction from G_0-arrested chicken immature erythrocytes was found to be enriched in active gene chromatin fragments (Delcuve and Davie, 1989).

Replacement variants H2A.Z and H3.3 are enriched in this fraction (Ridsdale and Davie, 1987). This is probably accounted for by the fact that newly synthesized core histones (H2A, H2A.Z, H2B, H3.3 and H4) are preferentially accumulated in this active chromatin fraction and depleted in repressed gene chromatin (Hendzel and Davie, 1990). Newly synthesized histone H5 was incorporated into the chromatin in a random way (Hendzel and Davie, 1990).

Analysis of the histone distribution in *Tetrahymena* has shown that histone variants hv1 (H2A.F/Z homologue) and hv2 (replacement histone H3) are localized in the transcriptionally active macronucleus but not in the transcriptionally inactive micronucleus (Bannon and Gorovsky, 1984). Antibodies against hv1 stained macronuclei at all stages of the *Tetrahymena* life cycle. Micronuclei were stained with this serum only during early stages of conjugation, preceding the brief period of transcriptional activity in the micronuclei during meiotic prophase (Stargell et al., 1993). Anti-hv1 antibodies also reacted with mammalian nucleoli, presumably with the H2A.F/Z histone in the chromatin of the highly active ribosomal genes (Allis et al., 1982). Enrichment in the nucleolar chromatin was also observed for the H2A.X variant (Bhatnagar et al., 1984). In *Drosophila* histone H2A.2 (which was earlier designated as D2 and which probably corresponds to H2AvD, the H2A.F/Z homologue) is preferentially associated with the interbands of the polytene chromosomes, which have a less compacted structure than bands (Donahue et al., 1986). Taken together, the localization data suggest that histone H2A replacement variants are preferentially localized in active chromatin and might have a function in the establishment or maintenance of a transcriptionally active/competent state. Their more or less constant level (~10%) in chromatin of most cell types is in agreement with this interpretation.

Histone H3.3, on the other hand, cannot be associated exclusively with the active chromatin, since it is the major histone H3 in certain cell types (see above). This does not contradict its preferential exchange and accumulation in active chromatin and might simply reflect a higher rate of nucleosome replacement in transcribed chromatin.

Histone $H1^0$ distribution between active and repressed genes was addressed in several studies. Fractionation of nucleosomes from adult mouse liver according to their $H1^0$ content has demonstrated that the transcriptionally inactive gene for α-fetoprotein, but not the active albumin gene, is preferentially associated with the $H1^0$ histone (Roche et al., 1985b). However, the α-fetoprotein gene does not become derepressed in the livers of mice, totally lacking $H1^0$ (Sirotkin et al., 1995). In a different study, involving immunofractionation of chromatin with anti-$H1^0$ antibodies, it was shown that while highly inducible genes may be segregated into $H1^0$-depleted chromatin regions, the $H1^0$ content is not necessarily different in transcribed and non-transcribed genes (Mendelson et al., 1986). Immunoelectron microscopy with antibodies against histone $H1^0$ demonstrated the presence of this histone in condensed chromatin areas, including perinucleolar chromatin, and in perichromatin regions, described as preferential sites of pre-mRNA synthesis (Gorka et al., 1993). This suggests that the histone $H1^0$ is not entirely excluded from transcriptonally active chromatin.

8.4. Differences in modification patterns

It is generally recognized that histone modifications, especially acetylation and phosphorylation have important functional consequences (Wu et al., 1986; Bradbury, 1992). Differential modification patterns of histone variants might, therefore, reflect functional differences.

In animals the replacement histones display the same spectrum of modifications as the replication dependent histones. For example, comparison of acetylation, mitogen-stimulated phosphorylation and methylation of replication-dependent variants H3.1 and H3.2 with the replacement H3.3 variant in mammalian cells revealed no major differences (Barratt et al., 1994; Annunziato et al., 1995). Histone H2A.Z undergoes acetylation and ubiquitinilation, as well as the replication-dependent H2A histones (Ridsdale and Davie, 1987; Nickel et al., 1987, Jasinskiene et al., 1995). Specific phosphorylation of the H2A.X variant in *Xenopus* egg extracts was reported to be necessary for the assemly of properly spaced nucleosomes on naked DNA (Kleinschmidt and Steinbeisser, 1991). However, subsequent studies of sperm genome decondensation in the same extracts have shown that both H2A.X and major H2A variant are phosphorylated and this phosphorylation is not essential for proper nucleosome spacing (Dimitrov et al., 1994).

In higher plants replacement histone H3.2 has a higher level of acetylation, than the cell-cycle regulated H3.1 variant (Waterborg et al., 1987; Waterborg, 1990; 1992). Acetylation occurs in the N-terminal domain, which is identical between H3.1 and H3.2 (Waterborg, 1990). The highly acetylated histone H3.2 fraction corresponds to the labile fraction of this histone variant (Waterborg, 1993). Quantitative and qualitative differences between the methylation patterns of the H3.1 and H3.2 N-termini were also observed (Waterborg, 1990). High acetylation level of histone H3.2 are compatible with the idea that this variant is preferentially accumulated in transcriptionally active chromatin, because there is a correlation between high levels of histone acetylation and transcriptional activation (Wolffe and Pruss, 1996c). The more apparent differences between the acetylation patterns of the H3 variants in plants, compared with animals may reflect a more important role for the histone H3 acetylation in plants, where the histone H3 N-terminus is the main target of dynamic acetylation (Waterborg, 1992). This correlates with an extremely high turnover rate of the replacement histone H3.

8.5. Genetical analyses of the function

8.5.1. Core histones

Functions of a particular gene can be most directly assessed by genetical means. Most extensive genetical studies of histone genes were performed in budding yeast, *Saccharomyces cerevisiae*. There are two genes for each of the core histones, which are organized in two H2A-H2B and two H3-H4 pairs. The two H2A genes, as well as the H2B genes, slightly differ from each other in the sequence of the encoded proteins. The histone H2A and H2B subtypes can

substitute for each other. They are dispensable as long as there remains one functional H2A and H2B gene (Grunstein et al., 1984). However, the two H2A-H2B pairs are not functionally equivalent, since only one of them can compensate for the absence of its partner and they have different functions in heat shock response (Norris and Osley, 1987).

Due to the fact that budding yeast does not contain replacement histones, it cannot be directly used for analysis of their function. However, the excellent genetical tools, available in this species, can be used to compare the properties of replication-dependent and replacement histone variants from other organisms. It was shown, for example, that yeast histone H2A can be substituted for the *Tetrahymena* major H2A variants. The *Tetrahymena* homologue of the H2A.F/Z histone, the hv1 variant, could not substitute for the yeast H2A proteins (Liu et al., 1996). This experiment, although it demonstrates nicely the functional difference between the major H2A proteins and the hv1 variant from *Tetrahymena*, provides no clues to the nature of this difference, since a large variety of factors might contribute to a failure of a particular protein to work in a heterologous system.

A histone H2A.F/Z-like protein was identified in the fission yeast, *S. pombe* (Carr et al., 1994). Deletion of the gene *pht1*, encoding this protein, is not lethal, but it causes slow growth, altered colony morphology, increased resistance to heat shock and minichromosome missegregation (Carr et al., 1994). This suggests important, though not essential functions of this protein. It should be noted that at present it is unclear whether these functions are different from those of other H2A histones, since there is no data on deletion of major histone genes in fission yeast.

Genetic manipulation in *Tetrahymena* has shown that the hv1 histone is essential in this species and cannot be compensated by the genes, encoding major H2A histones (Liu et al., 1996). In the same study it was shown that either of the two major H2A genes can be completely knocked out without effect on vegetative growth. However, at least one major H2A gene is essential, which means that the hv1-encoding gene cannot be the sole source of the H2A protein. Since the major H2As differ from the hv1 variant both in their primary structure and mode of regulation, it is unclear whether either of these features alone or both of them are essential.

In *Drosophila*, a deletion in the H2AvD gene is homozygous lethal, which indicates that this histone variant performs an essential function (van Daal and Elgin, 1992). Also in this case it is not completely clear whether the phenotype results from the absence of this particular protein variant or from the lack of an H2A gene, regulated in a specific way. The effect of the H2A dosage at a particular stage cannot be excluded, since the mutant flies die as larvae, when the synthesis of the cell-cycle regulated histones is reduced.

An obvious way to distinguish between the effects of protein structure and the mode of regulation on the replacement histone function is to try to rescue replacement histone mutants with synthetic genes, where the open reading frame

of a replacement histone is substituted with the frame of the corresponding cell-cycle regulated gene. One complication with performing such experiments can arise from the presence of regulatory sequences within the coding region itself (see above). It might prove to be difficult to separate coding and regulatory functions of a particular histone gene.

8.5.2. Linker histones

Linker histones have less crucial functions in the organization of the chromatin than core histones. They are apparenly absent in yeast, and, although present, are not essential in *Tetrahymena* (Shen et al., 1995). In vitro studies have shown that histone H1 is not required for the formation of condensed metaphase chromosomes in the *Xenopus* egg extracts (Ohsumi et al., 1993). This is consistent with the observation that in cell-free extracts of *Xenopus* eggs, linker histones are nonessential for the assembly of nuclei which look normal by cytological criteria and their ability to replicate DNA and import proteins (Dasso et al., 1994).

It is not surprising, therefore, that changes in the linker histone subtypes might have no grave consequences. It was shown that although histone $H1^{0}$ is an evolutionary ancient H1 replacement type which is regulated in a highly complex way, mice homozygous for a null mutation in this gene develop normally (Sirotkin et al., 1995). In such mutant mice the ratio of the total amount of linker histones to that of core histones was not measurably affected even in cells where $H1^{0}$ normally constitutes about 30% of total H1. Relative proportions of different somatic subtypes were not affected either. This indicates that the expression of all cell-cycle regulated histone H1 genes is stimulated in such mice, and together these genes compensate for the absense of $H1^{0}$. This observation supports the idea that if histone dosage can be adjusted, no severe phenotype results from the disruption of a particular histone gene. It also demonstrates that at least in this case the necessary basal level of expression can be provided by replication-dependent genes and that the level of histone of a particular type (in this case H1) can be "sensed" by all members of the multigene family. Similar observations were made when individual members of the H2B and H3 families were disrupted in cultured chicken cells (Takami et al., 1995a, b).

Effects of an increased dosage of histone $H1^{0}$ were tested in cultured mouse cells (Brown, D.T. et al., 1996). Cells, overproducing $H1^{0}$ have normal viability and no long-term growth arrest. However, these cells exhibit transient delay in cell cycle progression and reduced exression of a number of genes. In contrast, similar overexpression of another H1 variant, H1c, has no effect on cell cycle progression and leads to increased expression of some genes. These results suggest that the two linker histone variants are not functionally equivalent and support the idea about the involvement of $H1^{0}$ in gene repression. Selective transcriptional repression and proliferation arrest was also observed as a result of histone H5 overexpression in cultured cells (Sun et al., 1989).

Evidence for functional difference between linker histone subtypes was also obtained in the studies of early development in *Xenopus*. Cleavage linker his-

tone H1M (B4) is substituted by the somatic histones H1 in the course of embryogenesis (Dimitrov et al., 1993). Both in vitro and in vivo data indicate that the incorporation of the somatic histone H1 leads to transcriptional repression of certain genes, such as the oocyte 5S rRNA genes (Wolffe, 1989; Chipev and Wolffe. 1992; Bouvet et al., 1994). An in vivo decrease in histone H1 within chromatin facilitates the activation of the oocyte 5S rRNA genes, but has little impact on the somatic 5S rRNA genes (Bouvet et al., 1994). In these experiments histone H1 concentration was altered by microinjection of histone H1-encoding mRNA or a ribozyme, targeted against this mRNA, into *Xenopus* eggs. Such manipulations are comparable to genetic transformation experiments in other systems.

The data presented in this review show that many questions concerning the function of histone replacement variants remain unanswered. Elucidation of these questions is of general importance, since it is becoming increasingly clear that the chromatin structure represents a crucial factor in all aspects of the functioning of the eukaryotic genome.

Chapter II

Structure and expression of histone H3.3 genes in *Drosophila melanogaster* and *Drosophila hydei*

Anna Akhmanova[1], Petra C.T. Bindels[1], Jie Xu[1,2], Koos Miedema[1], Hannie Kremer[1,3] and Wolfgang Hennig[1,2]

[1]Department of Molecular and Developmental Genetics, Faculty of Sciences, Catholic University of Nijmegen, Toernooiveld, 6525 ED Nijmegen, The Netherlands

[2]Max-Plank-Guest Laboratory at the Institute of Cell Biology, Academia Sinica, Shanghai, China

[3]Present address: Department of Human Genetics, Catholic University of Nijmegen, 6525 Ed Nijmegen, The Netherlands

This chapter is published in:
Genome 38, 586-600 (1995)

Structure and expression of histone H3.3 genes in *Drosophila melanogaster* and *Drosophila hydei*

Anna S. Akhmanova, Petra C.T. Bindels, Jie Xu, Koos Miedema, Hannie Kremer, and Wolfgang Hennig

Abstract We demonstrate that in *Drosophila melanogaster* the histone H3 3 replacement variant is encoded by two genes, *H3 3A* and *H3 3B* We have isolated cDNA clones for *H3 3A* and cDNA and genomic clones for *H3 3B* The genes encode exactly the same protein but are widely divergent in their untranslated regions (UTR) Both genes are expressed in embryos and adults, they are expressed in the gonads as well as in somatic tissues of the flies However, only one of them, *H3 3A*, shows strong testes expression The 3′ UTR of the *H3 3A* gene is relatively short (~250 nucleotides (nt)) *H3 3B* transcripts can be processed at several polyadenylation sites, the longest with a 3′ UTR of more than 1500 nt The 3′ processing sites, preferentially used in the gonads and somatic tissues, are different We have also isolated the *Drosophila hydei* homologues of the two H3 3 genes They are quite similar to the *D melanogaster* genes in their expression patterns However, in contrast to their vertebrate counterparts, which are highly conserved in their noncoding regions, the *Drosophila* genes display only limited sequence similarity in these regions

Key words H3 3 histone variant, Drosophila, sequence comparison, alternative polyadenylation, testis expression

Résumé Les auteurs démontrent que chez le *Drosophila melanogaster* la variante d'histone H3 3 est codée par deux gènes, *H3 3A* et *H3 3B* Des clones d'ADNc ont été isolés pour *H3 3A* et *H3 3B* de même que des clones génomiques pour *H3 3B* Ces gènes codent pour des protéines identiques mais ils diffèrent au niveau de leurs régions non-traduites Les deux gènes sont exprimés dans les embryons et chez les adultes, ils s'expriment dans le gonades de même que dans les tissus somatiques Cependant, l'un deux, *H3 3A*, est fortement exprimé dans les testicules La région 3′ non traduite (3′ UTR) du gène *H3 3A* est relativement courte (~250 nucléotides (nt)) Les transcrits du gène *H3 3B* peuvent être coupés et polyadénylés à plusieurs sites dont le plus distant produit une 3′ UTR de plus de 1500 nt Les sites employés dans les gonades diffèrent de ceux utilisés dans les tissus somatiques Des homologues des deux gènes H3 3 ont également été isolés du *Drosophila hydei* L'expression de ces gènes est très semblable à celle observée chez le *D melanogaster* Cependant, contrairement à leurs homologues chez les vertebrés qui sont hautement conservés au niveau de leurs régions non-traduites, les séquences de ces genes chez le *Drosophila* montrent une similitude limitée dans ces regions

Mots clés variante d'histone H3 3, *Drosophila*, comparaison de sequences, sites de polyadénylation alternatifs, expression testiculaire

[Traduit par la Rédaction]

Introduction

Histones in higher eukaryotes are divided into two classes, replication-dependent and replacement (or basal) variants The synthesis of replication-dependent histones is coupled to the S phase of the cell cycle Their messenger RNAs are not polyadenylated, their genes contain no introns, and usually they are members of multicopy gene families (Schumperli 1986) In *Drosophila melanogaster* replication dependent histone genes are organized in an array of tandem repeats, each repeat bearing a copy of a histone gene of each type (H1, H2A, H2B, H3, and H4) (Matsuo and

Corresponding Editor P B Moens
Received November 9, 1994 Accepted January 9, 1995

A.S. Akhmanova, P.C.T. Bindels, J. Xu, K. Miedema, H. Kremer,[1] and W. Hennig.[2] Department of Molecular and Developmental Genetics, Faculty of Sciences, Catholic University of Nijmegen, Toernooiveld, 6525 ED Nijmegen, The Netherlands
J. Xu and W. Hennig. Max Plank Guest Laboratory at the Institute of Cell Biology, Academia Sinica, Shanghai, China

[1] Present address Department of Human Genetics, Catholic University of Nijmegen, 6525 ED Nijmegen, The Netherlands
[2] Author to whom all correspondence should be sent at the following address Department of Molecular and Developmental Genetics, Faculty of Sciences, Catholic University of Nijmegen Toernooiveld 6525 ED Nijmegen The Netherlands

Yamazaki 1989) The same organization is found in *Drosophila hydei* (Kremer and Hennig 1990)

Histone replacement variants are usually encoded by single copy genes, which contain introns and produce polyadenylated mRNAs (Schumperli 1986) Replacement histone variants also differ from their cell-cycle regulated counterparts at the level of amino acid sequence Histone H3 replacement variants have been described in *Tetrahymena* (Thatcher et al 1994), plants (Chaubet et al 1992), invertebrates (Fretzin et al 1991, Swenson et al 1987), and vertebrates (Brush et al 1985, Wells and Kedes 1985, Wellman et al 1987) In all these cases the replacement H3 variants differ by several amino acid substitutions from the corresponding replication dependent H3 histone The functional importance of these amino acid substitutions is unclear Phylogenetic analysis indicates that the replacement H3 variants, *h3 2* of *Tetrahymena*, plant H3 III, and animal H3 3 evolved independently (Thatcher et al 1994) It has been argued therefore that the replication independence of gene expression, rather than a particular protein structure, causes the existence of the H3 replacement variant The fact that there are no apparent differences in the post-translational modifications (acetylation and phosphorylation) between primary structure variants of the histone H3 (Barratt et al 1994) supports this interpretation

However H3 3 histone exemplifies one of the most conserved proteins in the animal kingdom, being identical in clam (Swenson et al 1987), *Drosophila* (Fretzin et al 1991), chicken (Brush et al 1985), mouse (Wellman et al 1987), and man (Wells and Kedes 1985) This argues strongly for the functional importance of the primary structure of this polypeptide, rather than simply for the regulatory aspects of its expression

In order to analyse the function and regulation of the expression of the H3 3 histone variant it is appropriate to study it in a genetically well characterized organism, such as *Drosophila* While our work was in progress (Kremer 1991), a *D melanogaster* cDNA fragment encoding histone H3 3 was described (Fretzin et al 1991) We have isolated and characterized the entire *D melanogaster* histone H3 3 transcription unit We also found a second H3 3 gene encoding exactly the same polypeptide, but completely divergent in its noncoding sequences Therefore, the genomic organization of histone H3 3 genes in *Drosophila* seems to be very similar to that of vertebrates, where the expression of two histone H3 3 genes, considerably divergent in their nucleotide sequence, has been reported (Brush et al 1985 Hraba Renevey and Kress 1989)

We also isolated the homologues of both genes from *D hydei* because an evolutionary approach can be very useful in identifying putative cis-acting elements that are important for gene expression Here we present the comparison of the sequences and a description of the expression patterns of the two H3 3 genes in both species, *D melanogaster* and *D hydei*

Materials and methods

Fly strains

The wild type strains of both *D melanogaster* Canton S and *D hydei* (Tubingen) were from our laboratory collection

Screening of the genomic and cDNA libraries

The following libraries were screened (*i*) a *D melanogaster* genomic library with partial *Eco*RI fragments in λDASHII, (*ii*) a *D hydei* genomic library with partial *Eco*RI fragments in λEMBL4, (*iii*) a *D hydei* testes poly(A)$^+$-cDNA library, (*iv*) a *D melanogaster* testes poly(A)$^+$-cDNA library, and (*v*) a *D melanogaster* embryonic (0–6 h) library in λZAPII The first three libraries were constructed in our laboratory, while the fourth was kindly provided by Dr T Hazelrigg and the fifth by Drs M Noll and E Hafen Phage plaques were lifted onto Hybond nylon filters (Amersham) and hybridization was carried out according to the protocols of the manufacturer The probes were labelled with [α-^{32}P]dCTP by nick translation (Sambrook et al 1989)

Recombinant DNA analysis

Plaque purified recombinant phages were amplified by liquid cultures and phage DNA was prepared as described by Sambrook et al (1989) DNA restriction fragments were subcloned into pBluescript II KS + or M13mp18 or M13mp19 The DNA sequences were determined by dideoxynucleotide sequencing of recombinant plasmids and of M13mp18 and M13mp19 subclones (Sambrook et al 1989) EMBL accession numbers are *D melanogaster* histone *H3 3A* embryonic cDNA (clone DmC6B), accession Number X82257, *D melanogaster* histone *H3 3B* testes cDNA (clone Dm37 5), accession Number X81205, *D melanogaster* genomic clone, encoding histone *H3 3B* and oligosaccharyltransferase, 48 kilodalton (kDa) subunit (from λDmE7 2), accession Number X81207, *D hydei* genomic clone, encoding histone *H3 3B* (from λDhE32 1), accession Number X81208, and *D hydei* histone *H3 3A* cDNA, reconstructed from several testes cDNA clones, accession Number X81206

Genomic DNA isolation and Southern blotting

Genomic DNA of both *D melanogaster* and *D hydei* was prepared from flies as described earlier (Huijser and Hennig 1987) Approximately 4 μg of DNA per lane was digested by a 3-fold excess of enzyme for several hours, electrophoresed on a 0 8% agarose gel, and transferred to Hybond nylon membranes

Southern blots were hybridized in 5× Denhardt's solution (1× Denhardt's solution 0 1% w/v polyvinylpyrrolidone, 0 1% w/v BSA, and 0 1% w/v Ficoll), 2× SSC (1× SSC 0 15 M NaCl plus 0 15 M sodium citrate), and 0 1% SDS at 65–68°C Membranes were washed in 2× SSC – 0 1% SDS at 50 or 68°C DNA probes were made by nick translation with [α-^{32}P]dCTP

RNA isolation and Northern blotting

Total RNA was prepared from different samples by the guanidinium–thiocyanate method (Chirgwin et al 1979) Embryos were collected from 0 to 18 h postfertilization, washed with tap water, washed with 0 9% NaCl, treated with 1% Triton X-100 in 0 9% NaCl, washed again in 0 9% NaCl, resuspended in 3% $NaClO_3$ for 3 min, washed in 0 9% NaCl, and frozen in liquid nitrogen Testes and ovaries were prepared from adult flies in a testis isolation buffer of 183 mM KCl, 47 mM NaCl, and 10 mM Tris–HCl (pH 7 4), during preparation these tissues, as well as the

carcasses, were frozen immediately in liquid nitrogen. Approximately 20 μg of total RNA per lane was electrophoresed on 1–1.2% agarose–formaldehyde gels (Sambrook et al. 1989) and transferred to Hybond nylon membranes.

Northern blots were hybridized in 0.5 M phosphate buffer, 7% SDS, 1 mM EDTA, and 1% BSA overnight at 60°C. Membranes were washed in 0.3 M phosphate buffer, and 1% SDS at 50°C (nonstringent conditions) or in 40 mM phosphate buffer, and 1% SDS at 60°C (stringent conditions). DNA probes were prepared as described above. The following probes were used (see also Fig. 1, 1C and 2C). (*i*) The *D. melanogaster H3.3B* gene (subclones from λDmE7.2) OST48, nucleotides 1–1107; INT, nucleotides 2750–3306; M3'A, nucleotides 3945–4143; M3'B, nucleotides 4341–4641; and M3'C, nucleotides 5029–5514. (*ii*) The *D. melanogaster* histone H3.2 variant probe is the *Eco*RI–*Sau*3A subclone from the *Eco*RI 5 kilobase pair (kbp) histone gene cluster repeat (corresponds to positions 2973–3838 of the sequence published by Matsuo and Yamasaki (1989)). (*iii*) The *D. melanogaster* histone *H3.3A* probe is the *Bam*HI–*Eco*RI cDNA fragment obtained by RT–PCR (reverse transcriptase polymerase chain reaction) on the RNA from ovaries with the primers oligo(dT) and H3-IV (see below), the positions in the sequence are 458–720. (*iv*) The *D hydei H3 3B* probes (subclones from λDhE32.1) HORF, nucleotides 2539–3427; H3'A, nucleotides 3191–3427; and H3'B, nucleotides 3428–4092. (*v*) The *D. hydei H3.3A* 3' UTR probe is the 3' terminal *Sal*I–*Sau*3A fragment of the RT–PCR product obtained on the RNA from ovaries with primers oligo(dT) and H3-IV (see below); the positions in the sequence are 630–786.

Cloning of the mDh3VT2 cDNA fragment

Total RNA (2 μg) from *D. hydei* testes was reverse transcribed in the presence of oligo(dT) with the help of Moloney Murine Leukemia Virus (M–MLV) reverse transcriptase (GIBCO BRL). PCR was carried out on the resulting cDNA with the primers H3-I and H3-II, the sequences of which were derived from the *D. hydei* cell-cycle regulated histone H3 ORF (open reading frame): H3-I, 5'-GCAGGCGGCCGCGCAGTTGGCCACTAAGGCAGC; H3-II, 5'-GCAGGCGCCGCAATGGCACACAAGTTTGTATC. The resulting PCR product was cloned and sequenced.

Rapid amplification of the cDNA ends (RACE)

This procedure was carried out essentially as described by Frohman (1990).

Amplification of the cDNA 5' ends

Total RNA from adult flies was prepared as described above. Poly(A)$^+$-RNA was isolated from total RNA using an oligo(dT)–cellulose column according to the protocol of the manufacturer (Stratagene). Synthesis and purification of the total oligo(dT)-primed cDNA were carried out with the aid of the Strategene cDNA synthesis kit. Oligo(dC) tails were added with terminal transferase (Boehringer Mannheim) according to the supplier's protocol. PCR was performed as described by Frohman (1990) with one gene-specific primer and an oligo(dG)-adapter primer with the sequence 5'-AGCTCTAGAGCGGCCGCAAGCTT(G)$_{12}$. The *D. hydei H3.3B* gene-specific primers were H3-III, 5'-CAGGATCCTGCTTGGTACGTGCCAT (positions 2570–2597 of the *D. hydei* histone *H3.3B* gene sequence), and Hh5', 5'-GATGATCACGGAGAACACGATCTGC (positions 338–320 of the same sequence) (see Fig. 1, 2C).

Amplification of the cDNA 3' ends

Total RNA from *D. melanogaster* or *D. hydei* ovaries was primed with oligo(dT), and cDNA was synthesized with M–MLV reverse transcriptase (GIBCO BRL) according to the suggestions of the supplier in the presence of 1 U/μL RNasin (Promega). The cDNA was used as a substrate for the PCR reaction with a gene-specific primer and an oligo(dT)-adapter primer having the sequence 5'-GACTCGAGTCGACATCGA(T)$_{17}$. The *D melanogaster H3 3B* gene-specific primers were Mh3'-I, 5'-CATCTAGATTCGATCACTGTCGT (positions 3898–3915 of the *D melanogaster H3 3B* gene sequence), and Mh3'-II, 5'-CATCTAGAAGACCGAATCGAGTAC (positions 4624–4641 of the same sequence). The *D. hydei* primer H3-IV sequence was 5'-CAGGATCCATGCCCAAGGACATTCA (positions 3006–3022 of the *H3.3B* gene: this region of the H3 3 ORF is identical in *D. hydei H3.3A* and *H3 3B*) (see Fig 1, 1C and 2C).

In each case PCR products were cloned and sequenced and for each product several independent clones were analysed.

Primer extension and S1 nuclease mapping

All experiments were carried out using total RNA, prepared as described above, from ovaries or carcasses of either species. The primer extension assay was carried out essentially as described by Sambrook et al. (1989). In this procedure 0.1 ng of primer was annealed to 20 μg of total RNA. After precipitation, the primer was extended by M–MLV reverse transcriptase (50 U/μL) (GIBCO BRL) in the presence of the buffer supplied by the manufacturer; 0.5 mM each of dATP, dGTP, and dTTP; 0 5 μM dCTP, 0.5 mM DTT; 1 U/μL RNasin; and [α-^{32}P]dCTP (10 mCi/mL (1 Ci = 37 GBq); 1 μL / 10 μL of the reaction mixture). After 30 min of incubation at 37°C the dCTP concentration was increased to 0 5 mM and the incubation was continued for another 30 min. The samples were loaded on a 12% polyacrylamide sequencing gel. Sequencing reactions, primed with the same oligonucleotide on an appropriate M13 template, served as a reference. For mapping the 5' end of the *D. melanogaster H3.3B* gene the primer Mh5' was used, which has the sequence 5'-GATGATCACTTTCAACACGAGACGA (positions 1384–1365) (see Fig 1, 1C); for mapping the 5' end of the *D. hydei H3.3B* gene the primer Hh5' was used.

For the S1 nuclease mapping assay a randomly labeled single-stranded DNA fragment was prepared. This was achieved by annealing the same primer that was used for primer extension to the single-stranded DNA of a M13 subclone covering exon 1 (*Hin*dIII–*Hin*cII fragment (positions 1103–1970) in M13mp19 for *D melanogaster* and *Hin*dIII–*Pst*I fragment (positions 3–700) in M13mp19 for *D. hydei*). The primer was extended in the presence of [α-^{32}P]dCTP as described by Sambrook et al. (1989). After digesting with *Hin*dIII, the single-stranded probe was purified

by electrophoresis on a 1% alkaline agarose gel and S1 mapping was carried out as described by Sambrook et al (1989) using 100–400 U/mL S1 nuclease (Pharmacia) at 20°C. Products of S1 nuclease digestions were analysed the same way as the primer extension products.

As a control in both experiments we annealed *D melanogaster* specific primer (or probe) to *D hydei* RNA and vice versa. In control experiments no specific signals were observed.

In situ hybridization on polytene chromosomes

In situ hybridization was performed as described by Hennig et al. (1982), except that the DIG labeled DNA probe was synthesized with the aid of DIG DNA labeling kit (Boehringer Mannheim) and the hybrids were visualized using FITC (fluorescein isothiocyanate) labeled sheep anti DIG Fab fragments (Boehringer Mannheim).

Sequence analysis

DNA sequence analysis was carried out using programs of the Genetics Computer Group, University of Wisconsin (Devereux et al. 1987). The number of synonymous substitutions per site was calculated according to Nei and Gojobory (1986) and the standard errors were calculated according to Nei and Jin (1989).

Results

Cloning and sequencing of the H3.3 histone genes

Animal histone H3.3 variant differs from the cell cycle regulated histone H3 by four amino acid substitutions: Ala → Ser at position 31, Ala → Ser at position 87, Val → Ile at position 89 and Met → Gly at position 90 (Wells et al. 1986).

A fragment of *D hydei* H3.3 cDNA was obtained by RT-PCR using two H3 specific primers (H3 I and H3 II, Fig. 1 2C) on *D hydei* poly(A)$^+$ RNA (clone mDh3VT2 (Kremer 1991)). This cDNA fragment is a part of the histone H3 ORF from amino acid position 26–105. The encoded polypeptide contains all four amino acid substitutions which is characteristic for the histone H3.3 replacement variant.

This cDNA fragment was used to screen *D melanogaster* and *D hydei* genomic libraries. We obtained two identical genomic *D melanogaster* clones (λDmE7.2 and λDm8.1, only the first one was analysed) and one *D hydei* clone (λDhE3.1) hybridizing strongly to the probe and displaying a restriction map different from that expected for cell cycle regulated histone gene repeats. Restriction fragments hybridizing to the probe were sequenced.

With the same probe we also screened a *D melanogaster* testes cDNA library and identified, by sequencing, one positive clone (Dm37.5) as H3.3 cDNA. It is 30 nucleotides (nt) shorter in the 5′ UTR and 34 nt longer in the 3′ UTR than the published cDNA (Fretzin et al. 1991) but is otherwise identical to it except for one nucleotide (a G → T substitution in the 3′ UTR in position 598).

Organization of the H3.3 transcription units

Comparison of the sequence of the *D melanogaster* genomic clone with the cDNA sequences showed that they are identical in the overlapping regions except for the single substitution mentioned above (λDmE7.2, like our cDNA clone, also has a T in position 598). The genomic sequence, however, is interrupted by two introns: one, 1883 nt long in the 5′ UTR and the other, 139 nt long, in the ORF (Fig. 1, 1B). The position of the second intron is the same as the position of the third intron in the human H3.3 gene (Wells et al. 1987).

Primer extension was used to establish the transcription start site of the *D melanogaster* histone H3.3 gene (Fig. 2A). It was mapped 1 nt upstream from the 5′ end of the published cDNA sequence and confirmed that this cDNA was full length. This result was confirmed by S1 mapping (data not shown).

When sequencing the region upstream from the *D melanogaster* histone H3.3 exon 1, in the near vicinity we found the putative C-terminal end of another ORF, 993 nt long, with the same transcriptional orientation. The stop codon of the ORF is only 341 nt upstream from the H3.3 transcription start site (Fig. 1, 1B). This ORF is 64% identical, on the nucleic acid level, and 76% similar, on the protein level, to the C-terminus of the oligosaccharyltransferase 48 kDa subunit (OST48) gene in dog (amino acid positions 129–445) (Silberstein et al. 1992).

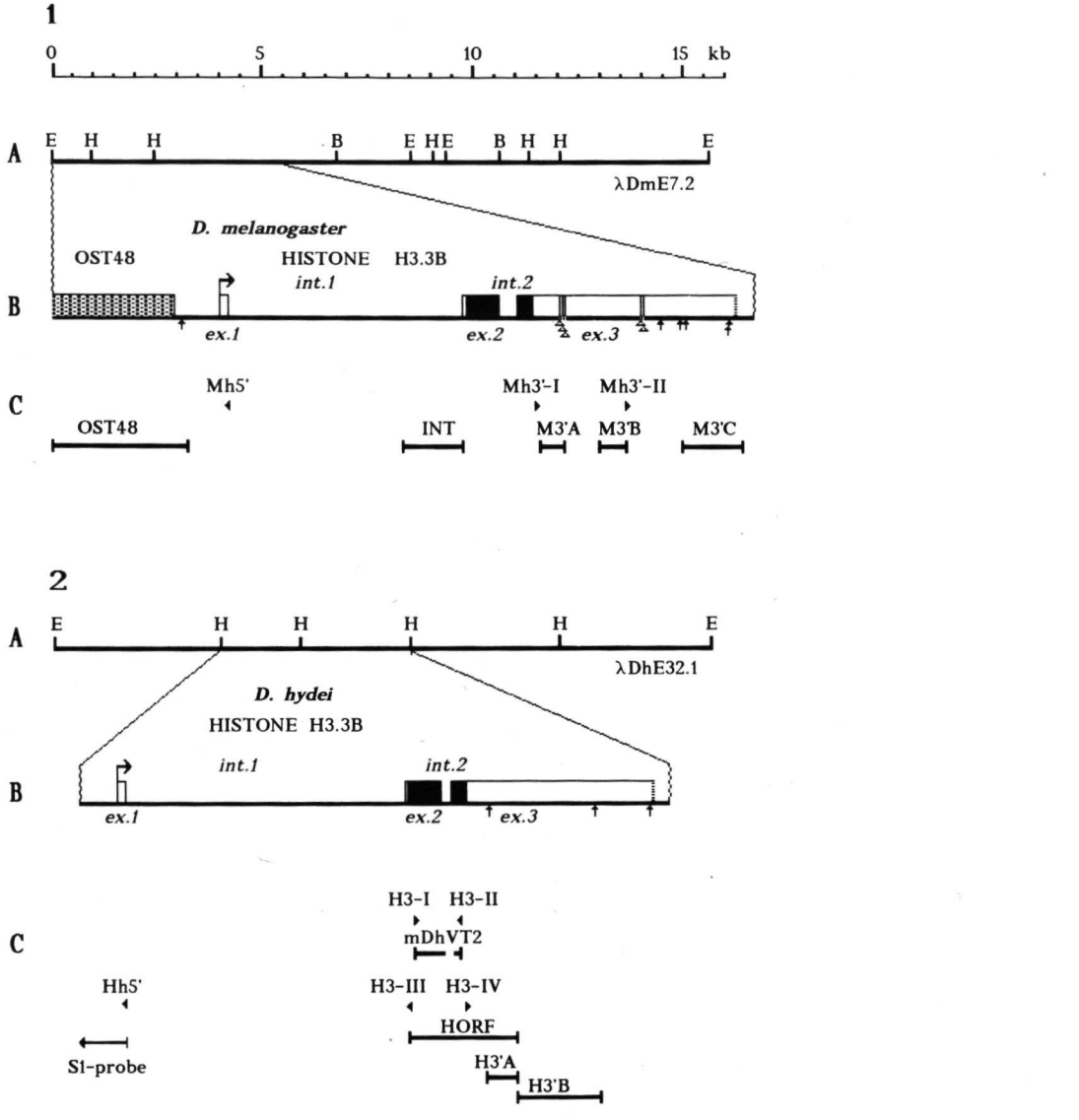
The *D hydei* H3.3 gene has a structure very similar to that of the *D melanogaster* H3.3 gene (Fig. 1, 2B). It also contains a small intron (58 nt) in the same position of the ORF and a large intron (2194 nt) in the 5′ UTR. To determine the transcription start site and the borders of the first intron we made use of the RACE technique (Frohman 1990), using primer H3-III, the sequence of which was derived from the start of the H3.3 ORF in clone λDhE32.1. The 5′ UTR of the H3.3 histone gene of *D hydei* was cloned in this way. The position of the transcription start site was confirmed by S1 mapping (Fig. 2B) and by another RACE experiment in which a 5′ UTR specific primer, Hh5-I, was used.

To investigate whether the close physical association between histone H3.3 and the OST48 homologous gene was conserved in *D hydei*, we hybridized the probe OST48 (Fig. 1, 1C) to λDhE32.1. This *D hydei* genomic clone contains 4 kbp upstream and 7 kbp downstream from the H3.3 transcription unit (Fig. 1, 2A). We observed no hybridization signal. In *D hydei*, therefore, the H3.3 and OST48 genes are separated by at least 4 kbp of DNA.

Expression of the H3.3 genes at the RNA level

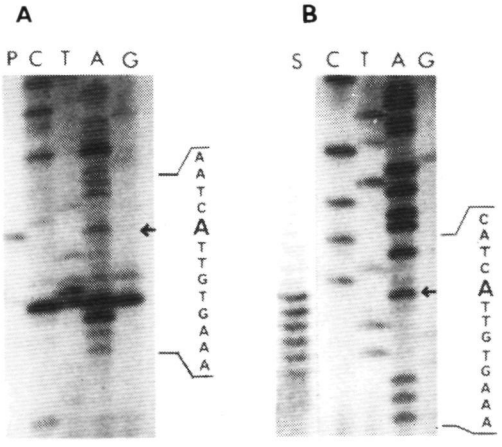
Northern blots with total RNA from *D melanogaster* male and female flies were probed under nonstringent conditions with a labeled insert of the cDNA clone Dm37.5 (Fig. 3A). Two major RNA species (~2400 and ~1700 nt) were found in the RNA of males, while additional bands of ~900 and ~550 nt were found in the RNA of females. The 550 nt band corresponds to the cell cycle regulated H3, as under stringent conditions it hybridized very strongly to the *D melanogaster* H3.2 probe (see Fig. 5C). It was not present on blots of poly(A)$^+$ RNA (not shown). Subsequently we hybridized the same blot with the M3′A probe (Fig. 3C), which contains only the proximal 3′ UTR sequence of the H3.3 gene (see Fig. 1, 1C). All the RNA species, except for the 550-nt transcripts, hybridized to the probe even under

Fig. 1. Structure of the *D. melanogaster* histone *H3.3B* genomic clone (1) and the *D. hydei* histone *H3.3B* genomic clone (2). (A) Restriction map. E, *Eco*RI; H, *Hin*dIII; B, *Bam*HI. (B) Structure of the sequenced part of the clone. The ORF of the oligosaccharyltransferase gene is shown by a shaded box; open boxes represent untranslated parts and black boxes represent translated parts of the exons of the histone H3.3 gene. Transcription starts are indicated by a horizontal arrow; polyadenylation signals are indicated by vertical black arrows. Open arrowheads indicate experimentally determined cleavage sites. (C) Positions of the primers and probes described in Materials and methods. Primers are shown by arrowheads; probes are shown by horizontal bars.



stringent conditions. To determine whether alternative 3′ processing was the reason for the different size classes of the transcripts, we subsequently probed a Northern blot of total RNA of females with two nonoverlapping distal 3′ probes, M3′B and M3′C (Fig. 3, D and E). As expected, the probe M3′B recognized only the two longer transcripts

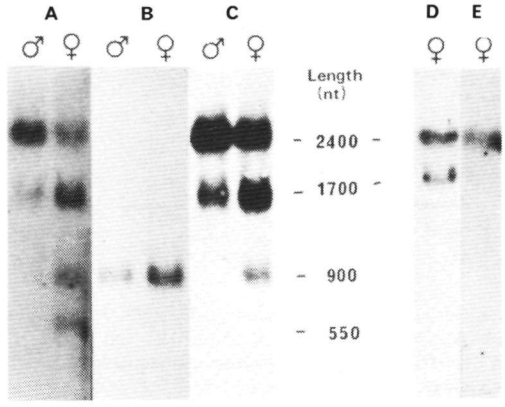
Fig. 2. Mapping of the transcription start sites of the *Drosophila H3.3B* genes. (A) Primer extension analysis of the 5′ terminus of *D. melanogaster* histone *H3.3B*. Primer Mh5′ was annealed to 10 μg of the RNA from ovaries and extended in the presence of reverse transcriptase (lane P). Lanes C, T, A and G show sequencing reactions of a M13 clone containing the exon 1 of the histone H3.3 gene initiated by the same primer. (B) S1 mapping of the 5′ terminus of *D. hydei* histone *H3.3B* gene. The labelled S1 probe was annealed to 10 μg of total RNA from ovaries and digested with 150 U/mL of S1 nuclease (lane S). Lanes C, T, A, and G show sequencing reactions of a M13 clone containing the exon 1 of the histone *H3.3B* gene initiated by the Hh5′ primer. The positions of the 5′ termini are indicated by arrows. The presence of additional weak bands in both experiments is most likely an experimental artifact. We cannot exclude, however, that there are additional, less frequently used, transcription start sites in both genes.



(2400 and 1700 nt), while M3′C hybridized only to the 2400-nt RNA species. The different H3.3 transcripts are therefore most likely the result of alternative polyadenylation.

Sequence analysis revealed the presence of two polyadenylation signals at positions 5418 and 5425. Utilization of these signals may account for the 2400-nt transcript. We did not try to localize this cleavage site more precisely. We attempted, however, to determine the two upstream processing sites using the RACE method. Total RNA from ovaries was used to synthesize cDNA with a $(dT)_{17}$ primer and PCR was performed on this cDNA with the oligo(dT)-adapter primer and one gene-specific primer (Mh3′-I or Mh3′-II) (Fig. 1, 1C), from which PCR products of ~250 and ~150 bp, respectively, were obtained, cloned, and sequenced. The positions of the polyadenylation sites obtained in this way are indicated in Fig. 4. It is noteworthy that there is no polyadenylation signal in the whole 3′ UTR of the smallest (900 nt) transcript. The 3′ processing sites of the 1700-nt transcript are preceded by a very TATA-rich sequence; however, it does not contain a conventional polyadenylation signal (AATAAA or

Fig. 3. Northern blots of *D. melanogaster* RNA. Hybridization with Dm37.5 cDNA (A), *H3.3A*-specific probe DmpE1 (B), probe M3′A (C), probe M3′B (D), and probe M3′C (E). The length of the transcripts was determined with the help of the GIBCO BRL RNA ladder. Washing conditions were nonstringent in A and stringent in B, C, D, and E.



ATTAAA). From the sequence analysis of the PCR products it appears that 3′ processing of the transcripts of both size classes occurs at several sites that are close to each other (Fig. 4).

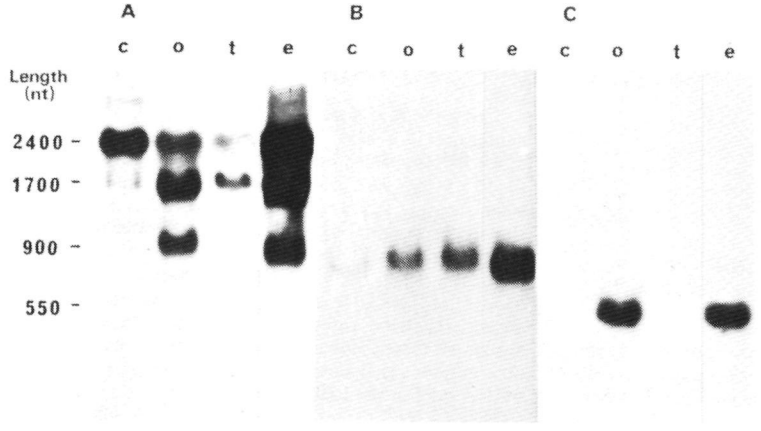
Using the M3′A probe we compared the histone H3.3 expression in ovaries, testes, carcasses (i.e., the remains of the fly after taking out the testes or ovaries), and embryos (0–18 h) (Fig. 5A). H3.3 mRNAs of all three size classes were strongly expressed in embryos; ovaries and testes showed preferential expression of the shorter transcripts, while in carcasses the largest transcript was the most abundant. When the same Northern blot was hybridized to histone H3.2 ORF probe (Fig. 5C), we obtained very strong signals on embyo and ovary RNA, but almost no signal on testes and carcass RNA. It seems that H3.3 messengers constitute the predominant part of the H3 mRNA pool in the latter tissues.

Histone H3.3 expression in *D. hydei* was analysed in a similar way. Hybridization of the probe HORF (containing full ORF, second intron, and part of the 3′ UTR) to the Northern blot under nonstringent conditions revealed transcripts of three size classes: 2100, 1000, and 550 nt (Fig. 6A). The 550 nt mRNA size class was assumed to represent the cell-cycle regulated histone H3 transcript and was therefore not analysed further. Hybridization with the distal (H3′B) and proximal (H3′A) probes revealed the largest transcript only (Fig. 6, B and C). Therefore, the *D. hydei* H3.3 gene that we have isolated cannot be the template for the 1000 nt transcripts.

Comparison of the *D. hydei* H3.3 gene expression in ovaries, testes, carcasses, and embryos, using probe H3′A (Fig. 6D) showed that the 2100-nt transcript is the most abundant in all RNA samples. Small amounts of RNA species 1800 nt long are present in ovaries (see arrow in Fig. 6D).

Fig. 4. Positions of the 3′ processing sites in the *D. melanogaster H3.3B* gene determined with the aid of the RACE technique. Arrowheads above the sequence indicate the ends of the PCR clones; the numbers above them denote the number of clones obtained and sequenced for each site. The ends of the cDNAs (from Fretzin et al. (1991) and from analysis of Dm37.5) are indicated by arrowheads below the sequence.

Fig. 5. Northern blots of *D. melanogaster* RNA. Hybridization with *H3.3B*-specific probe M3′A (A), *H3.3A*-specific probe DmpE1 (B), and H3.2-specific probe (C). Lane c, RNA from fly carcasses; lane o, RNA from ovaries; lane t, RNA from testes, and lane e, RNA from embryos (0–18 h). Transcript length was determined as in Fig. 3. All blots were washed stringently.



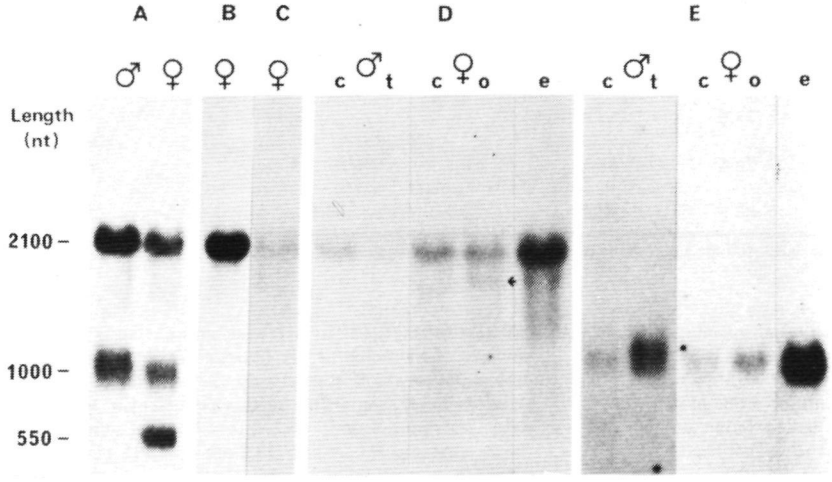
There is a second histone H3.3 gene in both *Drosophila* species

The origin of the *D. hydei* 1000-nt transcripts hybridizing to the H3.3 ORF remained unclear. Preliminary results showed that transcripts of this size class are very abundant in testis tissue. Therefore, a *D. hydei* testes cDNA library was screened with the mDh3VT2 probe and several truncated cDNA clones were obtained. All the clones apparently originated from the same gene, as they were identical in the overlapping regions. The reconstructed cDNA sequence contained a complete ORF encoding histone H3.3, ~100 nt of the 5′ UTR and ~280 nt of the 3′ UTR. The fact that this 3′ UTR is complete was confirmed with the aid of a RACE experiment in which an oligo(dT)-adapter primer and the primer H3-IV were used (see Materials and methods).

The probe containing exclusively the 3′ UTR sequences of this cDNA hybridized only to the 1000-nt transcript on a *D. hydei* Northern blot with RNA samples from embryos, ovaries, testes, and carcasses (Fig. 6E). The signal on testis RNA was very strong and the transcripts were more heterogeneous in length (starting from 1100 nt) than in other RNA samples. The reconstructed cDNA size (790 nt) is in good agreement with the transcript length on the Northern blot (taking into account the poly(A) tail of 200 nt).

We deduced that in *D. melanogaster* a second H3.3 gene must be present. To isolate this gene we performed a RT–PCR on RNA from *D. melanogaster* ovaries with the primers oligo(dT) and H3-IV. We obtained a cDNA fragment (clone DmpE1) 290 nt long, that contained the expected C-terminal part of the histone H3.3 ORF. The

Fig. 6. Northern blots of *D. hydei* RNA. Hybridization with probe HORF (A), probe H3′B (B), probe H3′A (C and D), and the histone *H3.3A*-specific 3′ UTR probe (E). Lane c, RNA from fly carcasses; lane o, RNA from ovaries; lane t, RNA from testes, and lane e, RNA from embryos (0–18 h). Transcript length was determined as in Fig. 3. A minor transcript of ~1800 nt is indicated by an arrow (D, lane o). Washing conditions were nonstringent in A and stringent in B, C, D, and E.



noncoding part of this cDNA, however, shared no similarity with the 3′ UTR sequences of the *D. melanogaster* H3.3 gene described above. The clone DmpE1 was used to screen a *D. melanogaster* embryonic library. One clone, DmC6B, was obtained. It contained a complete histone H3.3 ORF, ~90 nt of the 5′ UTR and the 3′ UTR sequence, identical to the clone DmpE1.

We used the clone DmpE1 to probe *D. melanogaster* Northern blots. The probe hybridized exclusively to a RNA species ~900 nt long (Figs. 3B and 5B), which is in good agreement with the length of the DmC6B cDNA (746 nt). The expression pattern of this H3.3 gene is similar to that of the two *D. hydei* genes that generate the 1000-nt transcript; it is relatively strongly transcribed in the testes.

Therefore, in both *Drosophila* species there are two different genes encoding histone H3.3. The *D. melanogaster* gene, represented by the cDNA clone DmC6B, is designated as histone *H3.3A* and the gene, represented by λDmE7.2, as *H3.3B*. Comparison of the lengths and expression patterns of the mRNAs suggests that the *D. hydei* H3.3 gene, which generates the 1000-nt transcripts, is the homologue of the *D. melanogaster H3.3A*. The *D. hydei* gene, represented by λDhE32.1, was likely to be homologous to the *D. melanogaster H3.3B*. Sequence comparisons between the genes confirm the proposed relationships (see below). This nomenclature is suggested by analogy with the mouse H3.3 genes. In mouse (and other vertebrates) also, two nonallelic H3.3 genes, A and B, were found. H3.3A is expressed as a single transcript of 1.2 kb, while three transcripts, ranging in size from 1.0 to 1.8 kb are synthesized from the H3.3B gene (Krimer et al. 1993). The actual relationship between the *Drosophila* and the vertebrate H3.3 histone genes is, however, unclear (see below).

Southern blot analysis and gene localization

The *D. melanogaster H3.3A* gene and both *D. hydei* H3.3 genes are single copy with respect to their 3′ UTRs (not shown). Also, the large intron of the *D. melanogaster H3.3B* gene is a single copy DNA segment (Figs. 7A and 7B; lane 1). However, the hybridization of the same genomic Southern blot with the two nonoverlapping 3′ UTR probes M3′A and M3′B under nonstringent conditions revealed several bands, one of which was the same with both probes (Figs. 7A and 7B; lanes 2 and 3. The additional band recognized by both 3′ UTR probes is indicated by an asterisk). It is likely, therefore, that other functional or nonfunctional copies of the *H3.3B* histone gene exist elsewhere in the genome of *D. melanogaster*. The *D. melanogaster H3.3B* gene isolated as the insert of λDm7.2 was mapped by in situ hybridization with the intron (INT) probe and was found in region 9D on the X chromosome. The *D. melanogaster H3.3A* gene was mapped with the DmpE1 probe to region 25C on chromosome 2.

Genomic blot hybridization with the OST48 probe showed that it is likely a single-copy segment (see Fig. 8).

Sequence comparison of the *Drosophila* H3.3 genes

Comparison of the coding regions

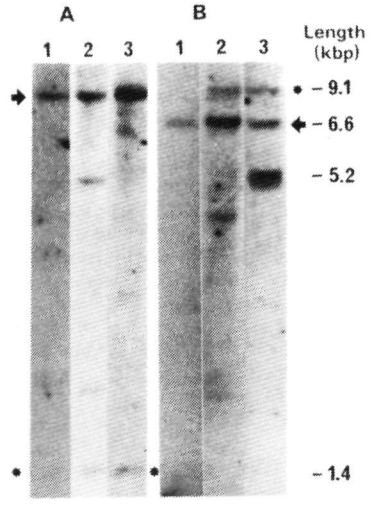
All the H3.3 ORFs from both *Drosophila* species encode exactly the same polypeptide. At the nucleic acid level, however, they are quite divergent (Table 1). While the H3.3B genes of the two species are rather similar, the similarity between the two H3.3A ORFs is lower than that of the *D. melanogaster H3.3A* and *H3.3B* genes compared with each other. Therefore, the comparison of the coding sequences alone is not sufficient to establish phylogenetic relationships between the *Drosophila* H3.3 genes.

Table 1. Percentage of similarity between the coding regions of the *Drosophila* histone H3 genes on the nucleic acid level.

	D. melanogaster H3.2	*D. hydei* H3.2	*D. melanogaster* H3.3A	*D. hydei* H3.3A	*D. melanogaster* H3.3B
D. hydei H3.2	83.7				
D. melanogaster H3.3A	80.3	78.1			
D.hydei H3.3A	80.2	78.6	83.0		
D. melanogaster H3.3B	78.3	76.6	86.1	80.1	
D. hydei H3.3B	80.0	80.5	82.7	83.2	86.4

Note: Percentage of similarity was calculated with the aid of the IFASTA program of the GCG sequence analysis program package.

Fig. 7. Southern blots of *D. melanogaster* genomic DNA. (A) *Eco*RI digests. (B) *Hin*dIII digests. Lane 1 is hybridized with the probe INT, lane 2 is hybridized with the probe M3′A, and lane 3 is hybridized with the probe M3′B. DNA fragments corresponding to the restriction map of λDmE7.2 are indicated by arrows; additional bands recognized by both M3′A and M3′B probes are indicated by asterisks.



We have computed the number of synonymous substitutions per site for the homologous pairs of the H3 genes (Table 2).[3] The number is somewhat lower for the variants than for cell-cycle regulated genes, and for *H3.3B* than for *H3.3A*. We compared these values with those obtained by Moriyama and Gojobory (1992), who used the same calculation method, in pairwise comparisons of genes between the species of the *Drosophila* and *Sophophora* subgenera (to which, respectively, *D. hydei* and *D. melanogaster* belong). According to this comparison, the *H3.3B* gene is a group I gene with a low level of synonymous substitution and *H3.3A* and H3.2 are group II genes with a moderate to high level of synonymous substitution.

Comparison of the base compositions at the synonymous positions of the codons revealed no major differences between cell-cycle regulated and replacement-variant genes (Table 2). G+C composition at synonymous sites is relatively low (with the exception of the *D. melanogaster H3.3A* it is ~45–65%), which places the genes in the group of low-biased genes (Shields et al. 1988). Also, the content of A in the third position of the fourfold degenerate sites is relatively high (Table 2); it is less than 10% for highly biased genes (Moriyama and Hartl 1993). The only apparent difference between the H3.2 and H3.3 genes with respect to codon usage is the number of different serine (Ser) codons (Table 2). Codons AGC(T) are avoided in the H3.3 genes, but are preferred in the H3.2 genes.

Comparison of the noncoding regions

Significant sequence similarities in noncoding regions could be found in pairwise comparisons of *H3.3A* and *H3.3B* genes from the two *Drosophila* species. When the *H3.3A* and *H3.3B* genes within one species were compared with each other, no such similarities could be found. This confirms our assumptions on the phylogenetic relationships between the histone H3.3 genes of the two *Drosophila* species, based on the transcript length and the expression patterns (see above).

Detailed comparison of the UTRs of the *D. hydei H3.3B* and *D. melanogaster H3.3B* genes revealed only limited stretches of similarity. Analysis of the promoter regions showed the absence of a TATA box in both species; a perfect match to the CAATT box (Dynan and Tjian 1985) was found in the *D. hydei* sequence (at position −72 from the transcription start) but not in the *D. melanogaster* sequence. The sequences in the position range −250 to −100 are very AT rich in both species (77% A+T). The

[3] Our value of K_s for cell-cycle regulated genes is higher than that obtained by Fitch and Strausbaugh (1993) because we used a different computation method.

Table 2. Base composition of synonymous positions the number of Ser codons and the number of synonymous substitutions per site in the coding regions of *Drosophila* histone H3 genes

	Base composition[a] (%)				$(G+C)_s^b$ (%)	Number of Ser codons		K_s^c (SE)[d]
	G	A	T	C		TCN	AG^C/T	
H3 2								1 50 (0 36)
D melanogaster	18 9	10 8	36 5	33 8	57 9	2	4	
D hydei	6 5	16 9	44 1	32 5	48 3	2	4	
H3 3A								1 39 (0 31)
D melanogaster	30 9	11 1	12 3	45 7	78 4	5	1	
D hydei	15 3	13 9	36 1	34 7	45 5	6	0	
H3 3B								1 00 (0 19)
D melanogaster	19 0	10 1	40 5	30 4	64 3	6	0	
D hydei	13 5	9 5	36 5	40 5	56 2	6	0	

[a] Base composition at the third positions of the fourfold degenerate codons
[b] G+C composition of all synonymous positions
[c] Number of synonymous substitutions per site computed according to Nei and Gojobory (1986)
[d] Standard error computed according to Nei and Jin (1989)

Fig. 8. Southern blots of *D melanogaster* genomic DNA hybridized with OST48 probe Lane 1 *Xba*I digest, lane 2 *Xho*I digest

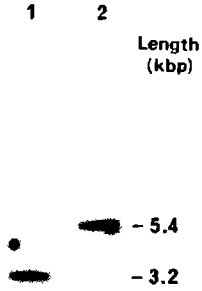


Fig. 9. Direct repeats in the promoter regions of the *Drosophila* replacement histone variants Nucleotide positions relative to the transcription start site (denoted +1) are indicated

D melanogaster	**-80**	CAATCGAGTG
H3 3 gene	**-70**	TAATCGAATA
	-60	CAATCGAAGT
D hydei	**-114**	TTATCGATAG
H3 3B gene	**-41**	CAGTCGATAC
	-16	TAATCGATAC
D melanogaster	**-84**	ATATCGATTG
H2AvD gene	**-69**	ACATCGATAT
	-1	TAGTCGAAAC

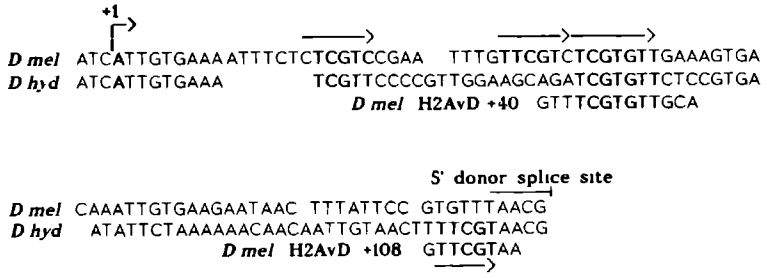
sequences from −100 to −1 are more GC rich (52 and 69% A+T in *D melanogaster* and *D hydei*, respectively) However no GC boxes (Dynan and Tjian 1985) could be found in either of the sequences They both contain three copies of a 8 bp repeat with the consensus AATCGAT(A)A(T) (see Fig 9) Two copies of a similar sequence (T(C)ATCGATT(A)) were found in the promoter region of *D melanogaster* H2A replacement variant *H2AvD* (van Daal et al 1990)

The 12 nt around the transcription start site are identical in *D melanogaster* and *D hydei* (Fig 10) Within exon 1 a conserved heptanucleotide was found that was also present within the *H2AvD* 5 UTR (van Daal et al 1990)

The splice junctions of the two introns in both species are in good agreement with the consensus (Mount et al 1992) Within the first intron, only short regions of similar sequences could be found The most prominent of these regions (with the exception of oligo(dT) and oligo(dA) stretches) are displayed in Fig 11 The second introns diverge completely

The 3 UTRs are also rather diverged In both species they are AT rich and contain multiple A stretches and $(TA)_n$ repeats Several boxes of similar sequences could be identified (Fig 11, $(TA)_n$ repeats are not shown) The longest similar sequence is one of 65 nt that is 96% identical (see Box 7 in Fig 11) We found no significant similarities between the 3 UTRs of *Drosophila* histone H3 3 genes and those of vertebrates Among invertebrates, H3 replacement variant cDNAs are characterized only in *Spisula solidissima* and *Tetrahymena thermophila* In the *Tetrahymena* H3 3 (*hv2*) gene the 3 UTR is AT rich and contains $(TA)_n$ and $(TAAA)_{2-3}$ repeats In this respect it

Fig. 10. Alignment of the sequences around the transcription start and of exon 1 between *D melanogaster* (*D mel*) and *D hydei* (*D hyd*) histone *H3 3B* genes The transcription start is marked with +1 Direct repeats, which are also present in the 5′ UTR of the H2AvD gene, are indicated by arrows



is similar to the histone *H3 3B* of *Drosophila* No significant similarities with the *Spisula* sequence were observed

Only the cDNA sequences of the *Drosophila H3 3A* genes are available at the moment The 5′ UTR sequences share no apparent similarities Also the 3′ UTRs are rather dissimilar (see Fig 12) Short boxes of similarity are present right after the stop codon and around the polyadenylation signal Both 3′ UTRs are rather GC rich and contain multiple short direct repeats They are shown as different types of arrows above and below the sequences in Fig 12

Discussion

Genomic organization and structure of the histone H3.3 genes in *Drosophila*

We have shown that there are two different functional histone H3 3 genes (*H3 3A* and *H3 3B*) in both species, *D melanogaster* and *D hydei* The genes encode exactly the same protein but they are completely divergent in their UTRs It seems unlikely that a third histone H3 3 gene exists in either of the species, because all the RNA bands hybridizing with the H3 3 ORF on the Northern blots are recognized by either of the two gene-specific 3′ UTR probes both in *D melanogaster* and *D hydei*

The histone H3 3 gene organization in *Drosophila* is thus very similar to that of vertebrates, where the H3 3 histone is also encoded by two different highly divergent genes (H3 3A and H3 3B) (Brush et al 1985, Hraba-Renevey and Kress 1989) The positions of the two introns in the *Drosophila H3 3B* gene are identical to those of the first and third introns in the human (Wells et al 1987) and chicken (Brush et al 1985) genes This argues for a monophyletic origin of the H3 3 histone replacement variants in the animal kingdom Also, as in vertebrates, the 3′ UTRs of the *Drosophila H3 3B* genes are very long (more than 1500 nt in *D melanogaster H3 3B*)

An interesting feature of the *D melanogaster* genomic clone described is the proximity of the histone H3 3B gene and OST48-like ORF We assume that OST48 is a functional housekeeping gene encoding an essential subunit of oligosaccharyltransferase, because of its high similarity to the corresponding sequence in dog (Silberstein et al 1992) Its ORF is not interrupted by stop codons and it behaves as a single-copy sequence on the Southern blot The localization of this ORF leaves only a maximum of 300 bp for the size of the *H3 3B* upstream regulatory region We cannot exclude, of course, that some regulatory elements of this gene are located within or upstream of the putative OST48-like gene However, more likely, some regulatory elements reside within the large first intron Many eukaryotic genes have a large intron within their 5′ UTR, in several cases such an intron has been shown to contain transcription enhancer and silencer elements (Gremke et al 1993, Gasch et al 1989) Our sequence alignment with the *D hydei* gene suggests some candidate regions for such elements Further experiments are required to establish their possible function

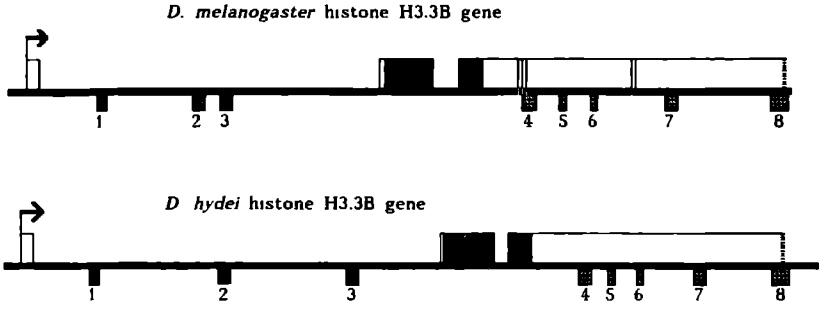
In mammals, H3 3 genes constitute a gene family with 20 to 30 members, most of which are, however, processed pseudogenes (Wells et al 1987) In *Drosophila*, three of the four isolated H3 3 genes are likely to be single copy Only for *D melanogaster* H3 3B do we find evidence that there might be a few (2–3) additional copies present in the genome

Evolution of the histone H3.3 genes in *Drosophila*

In *Drosophila*, genes that are expressed at a high level show relatively higher codon usage bias and lower substitution rates than genes expressed at a moderate or low level (Shields et al 1988, Moriyama and Gojobory 1992, Miedema et al 1994) The analysis of the *Drosophila* histone H3 3 coding regions revealed that they have a lower number of synonymous substitutions than do their cell cycle regulated counterparts The latter, however, are expressed much more strongly, at least in embryos This result is not unexpected, since it has been shown that the replication dependent histone genes in *Drosophila* display an unusually high synonymous substitution rate (Fitch and Strausbaugh 1993)

All the analysed genes display low codon usage bias It was suggested recently (Akashi 1994) that sequences more constrained at the amino acid level have higher codon bias, because the usage of preferred codons enhances translational accuracy This generalization apparently does not hold true for histones, which are completely conserved at the amino acid level It is possible that the multiplicity of the histone H3 encoding genes reduces the selection pressure on individual sites within the genes

Fig. 11. Relative positions and sequences of the most significant alignments in the first introns and 3′ UTRs of the *D. melanogaster* and *D. hydei* histone *H3.3B* genes. Gene structure is shown in the same way as it is in Fig 1 Blocks of similar sequence are numbered from 1 to 8. Sequence positions correspond to the numbering in the database entries. Polyadenylation signals are underlined.



Box 1. Dm 1739 TTTTCCAAAATGGCCGCCAAGCAGGAGAATAGCAAGAAATAAAAAAGTT
Dh 678 TTTTTCAAAATGGCGGCCA CTGCAGACAAGTTGGAAGTAAAAAGTTT

Box 2: Dm 22491 TTCAATGCGCGCCAAGTTTTACCGCTTGATTTTCCTTCCTTTTG TTTATG
Dh 1357 TTCAAATGGGCGCCAAGTTTTG TTTTCCTTCACATTGCTCTTATG

TTTACTCTTTTG
ATTATTGTTATG

Box 3 Dm 2394 AAATTGAAAACTACAACTCATTACTTCAGCGTGATTGATTC CGCATAAA
Dh 2040 AAACTAAAAACTACAACTCATTACTTCAACTTGATTGATTCTGCGCGCATAAA

GCGCGAAA AAAACTCAAGAGGCT
AAGCAAAACGCAAAATTCACGACGCT

Box 4: Dm 4112 ACTGTAATGCTTAAAAATTATTTTATAACTCTTATTATTACTATTATCAATTA
Dh 3310 ACTGTAATACTTA TTATTATTATTATTATATTTAGAAATGTATA

GTTTTTAAGTTTTTTTTCTCGCCTG
GTTTTTTAAGTTTGTAGATCGCCCG

Box 5: Dm 4296 CGTGTCGCC CTGCAACATCCGCCACCAACCACTCCAACA
Dh 3456 CGTGCCGCCATCTCGAACAT CGCCACCTA ACTCCAACA

Box 6. Dm 4465 GCCAATCT CGTCGGCTGGACGTGGCAAGGC
Dh 3614 GCCAAACGAAACCCTTCGGCGAGTGGGGGCAAGGC

Box 7: Dm 4866 TGGGCGTGACAGAGTTAGTTGATTAAATAAGATTTTAGATTTACATGCATTTT
Dh 3924 TGGGCGTG GTAGTTGATTAAATAAG TTATAGATTTACATGCATTTT

ATACGTTTAGGC
ATACGTTTAGGC

Box 8: Dm 5369 ACAGCAACAGCAAAAATAACTGGAAACATCAGGCAAAAAAAA AAAAGAA
Dh 4333 ATAGAAATCGTAAAAAAAAAACAAAACAAACAACAAAAAAAACCCCCAGACAA

AATAAAAAATAAACGGTTTTTTTCT ATAGAAGAAACCAAAAAAACCCAA
GAACTTAAATAAACGGTTTTTTCTAGAGAGAAAAAGCCAAAAAAACACAA

Codon biases are significantly different in the two *Drosophila* species analysed (Fitch and Strausbaugh 1993). This might explain why the two H3 3 variant genes in *D. melanogaster* are more similar than are the H3.3A genes between *D. melanogaster* and *D. hydei*. When the DNA sequences of the H3 3 ORFs are compared between *Drosophila* and individual vertebrate species, they are found to be 75–78% identical . The large number of substitutions and the influence of codon usage biases make it impossible to determine the phylogeny of histone H3.3 genes on the basis of the nucleic acid sequences of the coding regions. The highly divergent noncoding sequences are also of no help in this respect. It is therefore impossible to establish, in this way, a relationship between a particular *Drosophila* H3.3 gene and one of the vertebrate H3.3 genes.

The comparison of the *Drosophila* H3.3 genes with the vertebrate H3.3 genes reveals an interesting evolutionary

Fig. 12. Alignment of the 3′ UTR sequences of the *D melanogaster* and *D hydei* histone *H3 3A* cDNAs Arrows of different types above and below the sequences indicate direct repeats Sequence positions correspond to the numbering in the database entries

```
Dm TAAATGCCCATGTCCGCCATCTTGGATTGGAAAAGC        AGTG GCGGCAAAACTACC 559
Dh TAAATGGC ATGGTCGCCATCTTTAATTAGAAAAGCGTCGCAGTTCGTGTCGCTGCTAACC 563
   stop

Dm AACCC  ACAGCTGTCG CTTGGAGGGGGGAACTCGGAGCTGGAGGTGGAGGTGGAGAAG 616
Dh AACAAAGACAAC GTCGACGTCGCTGTT GAAC  GCATCTGCTGC GGCTGTTTACAAA 618

Dm CGGAAAA  CCAAACACC CGGC  CGCAGCTGCCG   CCGCTTAG          TACT 659
Dh CAACAAAGGCCGTCGACGTCGGCGTCGCAGCTGCTGTGTCTGCTTAGCCTTAAGTTTCAT 678

Dm TAGGTCTATCTATACTTACCG ATAAGTAATC    AATACACAATGT                 702
Dh AAGATTTAT TATACTTAAAATATAAGTAATTCATAATTCGTAATTTGTCGCATGCAA      735

Dm           TCTAATAAAATCGATTCATTTTTGTTT     GTAAATGTTTAACCAC      746
Dh AATTTGTCGCTCCAATAAAA    ATACATTTTTGTTTTCTGTGTAAAACTTAGATCC      789
```

aspect This is implied by the extensive conservation of the 5′ and 3′ UTRs of vertebrate H3 3 histone genes, which show a 85% similarity in the 520-nt long H3 3A 3′ UTR in chicken and man They are separated by more than 200 Ma (Hennig 1983) In *Drosophila* these genes apparently evolve much faster The divergence time between *D hydei* and *D melanogaster* is estimated to be 60 Ma (Throckmorton 1975), yet their histone H3 3 3′ UTRs share only limited stretches of similarity Such quick evolution of the untranslated regions does not reflect a general property of *Drosophila* genes For example, the 5′ and 3′ UTRs of the muscle-myosin heavy-chain gene are 93 and 96% identical, respectively, when compared in *D hydei* and *D melanogaster* (Miedema et al 1994)

In the generally divergent context of the noncoding DNA regions of the histone H3 3 genes, we found several blocks that are similar in sequence (see above) In the case of other genes, such conserved blocks have repeatedly been shown to be important for gene expression (see for example Michiels et al 1989) It will be of interest to test their function, especially for those elements that are also present in the *H2AvD* histone gene

H3.3 gene expression

It is apparent from Northern blot analysis (see above) that the H3 3 genes in *Drosophila* have their strongest expression in embryos This agrees with data on other replacement histone variants, such as *Drosophila H2AvD* (van Daal and Elgin 1992) and chicken H3 3 (Brush et al 1985), which are more strongly transcribed in embryonic than in adult tissues We performed transcript in situ hybridization of *D melanogaster* embryos with the two *H3 3B* 3′ UTR probes, M3′A and M3′B, the *H3 3A*-specific probe, and two probes specific for cell-cycle regulated H3 and H4 (not shown) As was expected, replication dependent histone probes gave stronger signals than the variant-specific ones At all stages rather uniform staining was observed with all five probes Therefore, whatever the function of the long 3′ UTR of the *Drosophila* histone *H3 3B* might be, it apparently does not direct cell-specific patterns of gene expression in early embryos

H3 3 histone mRNAs are also quite abundant in ovaries, most likely they are stored in *Drosophila* eggs together with the replication dependent histone mRNAs Storage of H3 3 mRNA in oocytes has been demonstrated earlier in clam (Swenson et al 1987)

Although H3 3 variant transcripts correlate to a certain extent, as to their abundance, with the cell cycle regulated histone messengers, the latter are subject to more strict regulation Their stability decreases when cells cease to divide (Anderson and Lengyel 1984) Therefore, they are hardly present in differentiated tissues such as are predominantly found in *Drosophila* adult carcasses H3 3 variant transcripts, in contrast, are readily detected there In vertebrates, H3 3 protein accumulates in nondividing adult tissues and can even become a predominant H3 subtype (Zweidler 1984) This quite likely also holds true for *Drosophila*

The abundance of H3 3 transcripts at a certain stage or in a specific tissue is not necessarily correlated with the level of synthesis of the corresponding protein For instance it was shown, for the rat replacement variant histone H1° that under certain conditions the amount of the protein decreases while its mRNA level increases in a cell proliferation dependent manner (Khochbin et al 1991) This possibly implies translational control of the expression of this histone variant In several systems, where translational control was demonstrated, it was mediated by direct sequence repeats within the 3′ UTR (Wickens and Takayama 1994, Ostareck-Lederer et al 1994) Therefore the multiple direct repeats found in the *Drosophila* H3 3A 3′ UTRs are of particular interest

The two *Drosophila* H3 3 genes are differentially expressed, as only one of them, *H3 3A*, displays a strong testis expression Remarkably, the testes transcripts are more heterogeneous in length than the carcass ones This

might be a result of differential cleavage, more likely, however, it is caused by additional polyadenylation in male germ line cells In *D melanogaster* a stage specific increase in the length of the poly(A) tail has been demonstrated for several testis-specific genes (Schafer et al 1990), this is apparently related to the translational control of the expression of these genes

Alternative 3'-end processing in the *Drosophila* H3.3 genes

The *D melanogaster* histone *H3 3B* transcripts can be processed at several alternative sites Some of the processing sites used are not preceded by a polyadenylation signal consensus sequence This is not surprising, since in many, if not all, cases where unconventional cleavage sites are used, alternative polyadenylation is involved (see Wahle and Keller 1992 and references therein) The positions of the processing sites of the two smaller transcripts were determined In both cases there are T-rich sequences located downstream from the cleavage sites (see Fig 4) Such downstream T rich elements have been shown to be important for correct 3' processing (Wahle and Keller 1992)

The resulting 3' UTRs differ in their length from ~200 to ~1500 nt, the shorter transcripts occur preferentially in ovaries, embryos, and testes This most likely results from differential polyadenylation rather than from differential stability of the transcripts Otherwise, one would not expect the longer messengers to accumulate in adult tissues It is known that destabilizing, rather than stabilizing, sequence elements control mRNA stability (Sachs 1993) It has been demonstrated in several systems that a longer messenger is likely to possess more such destabilizing elements than a shorter version and is likely to be more unstable (Brocard et al 1994, Sureau and Perbal 1994)

Alternative polyadenylation is quite common for eukaryotic genes (see Bond and Davidson 1986 and references therein) It is observed, for instance, in the mouse H3 3B and $H1^0$ genes (Hraba-Renevey and Kress 1989, Brocard et al 1994) In *D melanogaster* a complicated processing pattern is observed, for example, for the *c-abl* oncogene homologue, *Dash* (Telford et al 1985) Similar to what appears to be the case for the *H3 3B*, shorter transcripts are generated by maternal specific polyadenylation sites Such alternative polyadenylation, therefore, might reflect a general property of the 3' processing machinery in the germ cell line

The biological function of the alternative processing described is not clear especially as it is poorly conserved For the homologue of the *D melanogaster H3 3B* gene in *D hydei* we observed that the longest transcript is the most abundant one not only in carcasses, but also in ovaries and testes (low amounts of shorter transcripts are, however, found in ovaries) This correlates with the general lack of conservation of the 3' UTR sequences between these genes

Acknowledgements

We appreciate the generous gift of materials from Drs T Hazelrigg P Boender R Renkawitz Pohl, H Jackle, M Noll, E Hafen, and Ch Kirchhoff For technical support we are indebted to Wieli Jansen, Harry Harhangi, and Lisette Sanders We thank our colleages Harry Harhangi, Ron Hochstenbach, and Xiaoping Sun for valuable comments and discussions, as well as for help in conducting the experiments Financial support was provided by the Max-Planck-Gesellschaft zur Forderung der Wissenschaften and by the Deutsche Forschungsgemeinschaft

References

Akashi, H 1994 Synonymous codon usage in *Drosophila melanogaster* natural selection and translational accuracy Genetics, **136** 927–935

Anderson, K V, and Lengyel, J A 1984 Histone gene expression in *Drosophila* development multiple levels of gene regulation *In* Histone genes Structure, organization and regulation *Edited by* G S Stein, J L Stein, and W F Marzluff John Wiley and Sons, New York pp 135–161

Barratt, M , Hazzalin, C , Cano, E , and Mahadevan, L C 1994 Mitogen-stimulated phosphorylation of histone H3 is targeted to a small hyperacetylation-sensitive fraction Proc Natl Acad Sci U S A **91** 4781–4785

Bond, B J , and Davidson, N 1986 The *Drosophila melanogaster* actin 5C gene uses two transcription initiation sites and three polyadenylation sites to express multiple mRNA species Mol Cell Biol **6** 2080–2088

Brocard, M P, Rousseau, D , Lawrence, J J , and Khochbin, S 1994 Two mRNA species encoding the differentiation associated histone $H1^0$ are produced by alternative polyadenylation in mouse Eur J Biochem **221** 421–425

Brush, D , Dodgson, J B , Choi, O R , Stevens, P W , and Engel, J D 1985 Replacement variant histone genes contain intervening sequences Mol Cell Biol **5** 1307–1317

Chaubet, N , Clement, B , and Gigot, C 1992 Genes encoding a histone H3 3-like variant in *Arabidopsis* contain intervening sequences J Mol Biol **225** 569–574

Chirgwin, J M , Przybyla, A E , MacDonald, R J , and Rutter, W J 1979 Isolation of biologically active ribonucleic acid from sources enriched in ribonuclease Biochemistry, **18** 5294–5299

Devereux, J , Haeberli, P, and Marquess, P 1987 Introduction to the sequence analysis software package of the genetics computer group, University of Wisconsin, 1987 University of Wisconsin, Madison

Dynan, W S , and Tjian, R 1985 Control of eukaryotic messenger RNA synthesis by sequence-specific DNA-binding proteins Nature (London), **316** 774–778

Fitch, D H A , and Strausbaugh, L D 1993 Low codon bias and high rates of synonymous substitution in *Drosophila hydei* and *D melanogaster* histone genes Mol Biol Evol **10** 397–413

Fretzin, S , Allan, B D , van Daal, A , and Elgin, S C 1991 A *Drosophila melanogaster* H3 3 cDNA encodes a histone variant identical with the vertebrate H3 3 Gene, **107** 341–342

Frohman, M A 1990 RACE rapid amplification of cDNA ends *In* PCR protocols a guide to methods and applications *Edited by* M A Innis, D H Gelfand, J J Sninsky, and T J White Academic Press, San Diego pp 28–38

Gasch, A , Hinz, U , and Renkawitz-Pohl, R 1989 Intron and upstream sequences regulate expression of the *Drosophila* β3-tubulin gene in the visceral and somatic musculature, respectively Proc Natl Acad Sci U S A **86** 3215–3218

Gremke, L Lord, P C , Sabacan, L , Lin, S C , Wohlwill, A , and Storti R V 1993 Coordinate regulation of *Drosophila* tropomyosin gene expression is controlled by multiple

muscle-type-specific positive and negative enhancer elements Dev Biol **159** 513–527
Hennig, W 1983 Stammesgeschichte der Chordaten Verlag Paul Parey, Hamburg, Berlin
Hennig W, Vogt, P, Jacob, G, and Siegmund, I 1982 Nucleolus organizer regions in *Drosophila* species of the repleta group Chromosoma, **87** 279–292
Hraba-Renevey, S, and Kress, M 1989 Expression of a mouse replacement histone H3 3 gene with a highly conserved 3′ noncoding region during SV40- and polyoma-induced Go to S-phase transition Nucleic Acids Res **17** 2449–2461
Huijser, P and Hennig, W 1987 Ribosomal DNA-related sequences in a Y chromosomal lampbrush loop of *Drosophila hydei* Mol Gen Genet **206** 441–451
Khochbin, S, Gorka, C, and Lawrence, J-J 1991 Multiple control level governing H1(0) mRNA and protein accumulation FEBS Lett **283** 65–67
Kremer, H 1991 Spermatogenesis of *Drosophila hydei* chromatin reorganization and histone variants Ph D thesis, Catholic University of Nijmegen, Nijmegen, The Netherlands
Kremer, H, and Hennig, W 1990 Isolation and characterization of a *Drosophila hydei* histone DNA repeat unit Nucleic Acids Res **18** 1573–1580
Krimer, D B, Cheng, G, and Skoultchi 1993 Induction of H3 3 replacement histone mRNAs during the preco mitment period of murine erythroleukemia cell differentiat Nucleic Acids Res **21** 2873–2879
Matsuo, Y, and Yamazaki, T 1989 tRNA derived insertion element in histone gene repeating unit of *Drosophila melanogaster* Nucleic Acids Res **17** 225–238
Michiels, F, Gasch, A, Kaltschmidt, B, and Renkawitz Pohl, R 1989 A 14 bp promoter element directs the testis specificity of the *Drosophila* β2 tubulin gene EMBO J **8** 1559–1565
Miedema, K, Harhangi, H, Mentzel, S, Wilbrink, M, Akhmanova, A, Hooiveld, M, Bindels, P, and Hennig, W 1994 Interspecific sequence comparison of the muscle-myosin heavy chain genes from *Drosophila hydei* and *Drosophila melanogaster* J Mol Evol **39** 357–368
Moriyama, E N, and Gojobory, T 1992 Rates of synonymous substitution and base composition of nuclear genes in *Drosophila* Genetics, **130** 855–864
Moriyama, E N, and Hartl, D L 1993 Codon usage bias and base composition of nuclear genes in *Drosophila* Genetics, **134** 847–858
Mount, S M, Burks, C, Hertz, G, Stormo, G D, White, O, and Fields, C 1992 Splicing signals in *Drosophila* intron size, information content, and consensus sequences Nucleic Acids Res **16** 4255–4262
Nei, M, and Gojobory, T 1986 Simple methods for estimating the numbers of synonymous and nonsynonymous nucleotide substitutions Mol Biol Evol **3** 418–426
Nei, M, and Jin, L 1989 Variances of the average numbers of nucleotide substitutions within and between populations Mol Biol Evol **6** 290–300
Ostareck-Lederer, A, Ostareck, D H, Standart, N, and Thiele, B J 1994 Translation of 15-lipoxygenase messenger RNA is inhibited by a protein that binds to a repeated sequence in the 3′ untranslated region EMBO J **13** 1476–1481
Sambrook, J, Fritsch, E F, and Maniatis, T 1989 Molecular cloning a laboratory manual Cold Spring Harbor Laboratory, Cold Spring Harbor, New York
Sachs, A B 1993 Messenger RNA degradation in eukaryotes Cell, **74** 413–421
Schäfer, M, Kuhn, R, Bosse, F, and Schäfer, U 1990 A conserved element in the leader mediates post meiotic translation as well as cytoplasmic polyadenylation of a *Drosophila* spermatocyte mRNA EMBO J **9** 4519–4525
Schumperli, D 1986 Cell-cycle regulation of histone gene expression Cell, **45** 471–472
Shields, D C, Sharp, P M, Higgins, D G and Wright F 1988 "Silent" sites in *Drosophila* genes are not neutral evidence of selection among synonymous codons Mol Biol Evol **5** 704–716
Silberstein, S, Kelleher, D J, and Gilmore, R 1992 The 48 kDa subunit of the mammalian oligosaccharyltransferase complex is homologous to the essential yeast protein WBP1 J Biol Chem **267** 23 658 – 23 663
Sureau, A, and Perbal, B 1994 Several messenger RNAs with variable 3′ untranslated regions and different stability encode the human PR264/SC35 splicing factor Proc Natl Acad Sci U S A **91** 932–936
Swenson, K I, Borgese, N, Pietrini, G and Ruderman J V 1987 Three translationally regulated mRNAs are stored in the cytoplasm of clam oocytes Dev Biol **123** 10–16
Telford, J, Burckhardt, J, Butler, B, and Pirrotta, V 1985 Alternative processing and developmental control of the transcripts of the *Drosophila* abl oncogene homologue EMBO J **4** 2609–2615
Thatcher, T H, Macgaffey, J, Bowen J, Horowitz, S, Shapiro, D L, and Gorovsky, M A 1994 Independent evolutionary origin of histone H3 3-like variants of animals and *Tetrahymena* Nucleic Acids Res **22** 180–186
Throckmorton, L H 1975 The phylogeny ecology and geography of *Drosophila In* Handbook of genetics invertebrates of genetic interest Vol 3 *Edited by* R C King Plenum Press, New York pp 421–469
van Daal, A, and Elgin, S C 1992 A histone variant H2AvD, is essential in *Drosophila melanogaster* Mol Biol Cell, **3** 593–602
van Daal, A, White, E M, Elgin, S C, and Gorovsky, M A 1990 Conservation of intron position indicates separation of major and variant H2As is an early event in the evolution of eukaryotes J Mol Evol **30** 449–455
Wahle, E and Keller, W 1992 The biochemistry of 3′-end cleavage and polyadenylation of messenger RNA precursors Annu Rev Biochem **61** 419–440
Wellman, S E, Casano, P J, Pilch D R, Marzluff, W F, and Sittman, D B 1987 Characterization of mouse H3 3-like histone genes Gene, **59** 29–39
Wells, D, and Kedes, L 1985 Structure of a human histone cDNA evidence that basally expressed histone genes have intervening sequences and encode polyadenylated mRNAs Proc Natl Acad Sci U S A **82** 2834–2838
Wells, D, Bains, W, and Kedes, L)86 Codon usage in histone gene families of higher eukaryotes reflects functional rather than phylogenetic relationships J Mol Evol **23** 224–241
Wells, D, Hoffman, D and Kedes, L 1987 Unusual structure, evolutionary conservation of non-coding sequences and numerous pseudogenes characterize the human H3 3 histone multigene family Nucleic Acids Res **15** 2871 2889
Wickens, M, and Takayama, K 1994 Deviants — or emissaries Nature (London), **367** 17–18
Zweidler, A 1984 Core histone variants of the mouse primary structure and differential expression in histone genes *In* Histone genes Structure, organization and regulation *Edited by* G S Stein, J L Stein, and W F Marzluff John Wiley & Sons, New York pp 339–371

Chapter III

In vitro analysis of polyadenylation of the *Drosophila melanogaster* histone H3.3B pre-mRNA

Anna Akhmanova, Samuel I. Gunderson[1], Maria Polycarpou-Schwarz[1], Wolfgang Hennig and Iain W. Mattaj[1]

Department of Molecular and Developmental Genetics, Faculty of Sciences, Catholic University of Nijmegen, Toernooiveld, 6525 ED Nijmegen, The Netherlands

[1] European Molecular Biology Laboratory, Meyerhofstrasse 1, W-6900 Heidelberg, Germany

Abstract

Different types of transcripts are generated from the *Drosophila melanogaster* histone H3.3B gene in a tissue-specific way as a result of alternative polyadenylation. The activity of the three alternative processing sites from this gene was analysed in an in vitro cleavage and polyadenylation assay in the nuclear extract from HeLa cells. As expected, the efficiency of polyadenylation in vitro was dependent on the presence of the conventional polyadenylation signal, AAUAAA. An alternative processing site with a deviant hexamer, AAUAUA, present in several tandem copies, was also found to be quite active in the in vitro assay if a "precleaved" substrate was used.

1. Introduction

All animal mRNAs with the exception of those encoding replication-dependent histones bear a poly(A) tail at their 3' end. This poly(A) tail is added as a result of the 3' processing reaction necessary to generate mature mRNA (for review, see Wahle and Keller, 1992; Wahle, 1995; Keller, 1995). Transcription by the RNA polymerase II, which transcribes all protein-encoding genes, proceeds beyond the 3' processing site. The processing itself consists of cleavage and poly(A) addition reactions. Both are tightly linked in the cell, but can be uncoupled experimentally. Several sequence elements are necessary for correct processing. Most important is the highly conserved hexamer signal AAUAAA, localized 10-30 nt upstream of the polyadenylation site. Both cleavage and polyadenylation require this sequence element. Point mutations in this signal, with the exception of A->U substition in the second position, lead to a dramatic reduction in the 3' processing efficiency. In addition, a diffuse GU-rich or U-rich element downstream of the cleavage site, is required for efficient processing. Finally, sequences upstream of the poly(A) signal can influence its utilization. Such sequences are often found in viral systems, such as adenovirus, SV40, HIV and other retroviruses (see Prescott and Falck-Pedersen, 1994, and references therein). The biochemistry of the 3' processing is understood quite well, as correct cleavage and polyadenylation can be obtained in vitro, in nuclear extracts from HeLa cells (Moore and Sharp, 1984, 1985). Most of the essential components of this in vitro system were purified and characterized (for review, see Keller, 1995).

Many genes contain several polyadenylation signals in their 3' flanking region. These signals are often used to produce mRNAs with 3'UTRs of different length in a tissue-specific manner. Alternative polyadenylation is a way of regulating differential gene expression in certain systems (Alt et al., 1980; Nevins and Wilson, 1981; Helfman et al., 1986). However, the mechanisms determining the choice of a particular poly(A)-signal were investigated only in a few cases. They can involve U-rich sequences upstream of the AAUAAA signal (Prescott and Falck-Pedersen, 1994) or downstream polyadenylation enhancers (Lou et al., 1996). In *Drosophila*, female-specific polyadenylation of *dsx* transcripts is induced by the *tra-2* gene product, which binds to a series of repeats upstream of the polyadenylation site (Hedley and Maniatis, 1991). It is noteworthy, that in many cases of alternative

polyadenylation at least some of the processing sites are not preceded by a conventional polyadenylation signal (Wahle and Keller, 1992). It was suggested that additional factors are necessary for the recognition of these sites.

In course of our studies of the *Drosophila* histone H3.3 gene expression (Akhmanova et al., 1995) we have found that one of the two *D. melanogaster* histone H3.3-encoding genes, H3.3B, is subject to alternative 3' processing. The histone H3.3B mRNAs of different size classes are generated as a result of alternative utilization of three processing sites. No consensus polyadenylation signal is present in the vicinity of the two upstream poly(A) sites, which are preferentially used in the gonads and in the embryos. The histone H3.3B gene is highly expressed and the alternative polyadenylation in its 3'UTR can serve as a useful model system. As an initial step of such a study we have analysed the activity of the three alternative H3.3B processing sites in the in vitro polyadenylation assay in nuclear extracts from HeLa cells.

2. Materials and Methods

Table 1. Constructs for the in vitro polyadenylation assay

NAME	Positions in the H3.3B sequence	Primers used for PCR[1]	Sites for cloning in pBluescriptKS+	Linearization site
CON 1	4003-4275	AA1 & AA2	SacI, XbaI	XbaI
CON 2	4624-4865	Mh3'-II & AA3	XbaI, SpeI	SpeI
CON 3	5363-5546	AA4 & UP20[2]	SacI, HincII	HincII
CON 4	4003-4116	AA1 & AA5	SacI, SpeI	RsaI
CON 5	4624-4752	Mh3'-II & AA6	XbaI, SpeI	SpeI
CON 6	5370-5449	AA4 & AA7	SacI, XbaI	XmnI

[1] Sequences of the primers were as follows:
Mh3'-II 5'-CATCTAGAAGACCGAATCGAGTAC
AA1 5'-CAGAGCTCATTCATATTTACATGT
AA2 5'-CATCTAGATTCCCTCCTTCCAAAACG
AA3 5'-ACACTAGTCATCCCTTCGTTATGTG
AA4 5'-CAGAGCTCAGTGACAGCAACAGCA
AA5 5'-ACACTAGTACAGTAATGTTGTGTGTTG
AA6 5'-ACACTAGTTCTTTTCCTTCCGCA
AA7 5'-CATCTAGAATTTCTTCTATAGAAAAAAC
Plasmid p82SP1, containing the SalI-PstI fragment of the H3.3B gene (from position 3498 to 5666 (Akhmanova et al., 1995)) in pBluescriptKS+, was used as a template for PCR.
[2] UP20 - M13 -20 universal primer

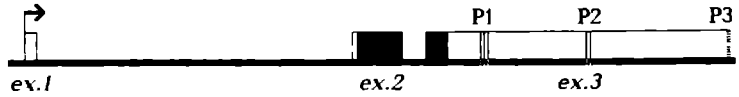
The generate the substrates for the in vitro cleavage and polyadenylation assay, different regions of the histone H3.3B gene were amplified by PCR and subcloned into pBluescriptKS+. The plasmids were linearized by restriction digestion and the labeled RNA substrates were synthesized by in vitro transcription with T7 RNA polymerase. The information about the sequence of the substrates is summarized in Table 1 and in Fig. 1. A substrate, containing the polyadenylation site of the U1 70K protein pre-mRNA was used as control. The in vitro polyadenylation assays were performed as described by Boelens et al. (1993), and the products of the assay were analysed on a denaturing polyacrylamide gel.

3. Results

Three types of transcripts with the length of 900 nt, 1700 nt and 2400 nt are generated from the *D. melanogaster* H3.3B gene (Akhmanova et al., 1995).. These transcripts result from alternative usage of three processing sites, which we will call here P1, P2 and P3 (see Fig. 1A). The exact polyadenylation sites of the two shorter transcripts (corresponding to P1 and P2) were determined by RT-PCR with oligo(dT) and H3.3B-specific primers. In both cases polyadenylation occurs at several adjacent positions, which are distributed over a 50 nt region in case of P1 and over 20 nt region in case of P2 (see Fig. 1B). The position of the P3 processing site was estimated from the length of the transcripts (2400 nt) on the Northern blots. The determination of the poly(A) addition sites of the 2400 nt transcript using the same strategy as for P1 and P2 was not possible due to the long oligo(A) stretches in the genomic region of the P3 site. However, it seems very likely that the polyadenylation signals AAUAAA at the position 5418 or 5425 are utilized to generate the 2400 nt transcript. First, their localization is in good agreement with the length of this transcript and there are no other poly(A) signals in this region. Most *Drosophila* genes contain a normal poly(A) signal upstream of the 3' processing site. Second, the sequences around these signals are conserved between the histone H3.3B genes of *D. melanogaster* and *D. hydei*. Third, experimentally determined polyadenylation sites of the histone H3.3A genes from both *D. melanogaster* and *D. hydei* are found in a sequence context, similar to that around the P3 processing site of the H3.3B gene (Akhmanova et al., 1995; see Fig. 2). We assumed that the polyadenylation at P3 would occur around position 5450, about 20 nt downstream of the poly(A) signals (as is usually the case; see Wahle and Keller, 1992).

DNA fragments containing the P1, P2 and P3 processing sites were used to prepare substrates for in vitro polyadenylation assay. Two types of constructs corresponding to each processing site were tested. The constructs of the first type (CON 1 for P1, CON 2 for P2 and CON 3 for P3) contained sequences both upstream and downsteam of the processing sites. These constructs were used to test both cleavage and polyadenylation. To test polyadenylation only, "precleaved" constructs were prepared. These constructs were made in such a way that they would end at one of the experimentally determined 3' processing sites (CON 4 for P1, CON 5 for P2 and CON 6 for P3; the end of CON 6 was chosen arbitrarily).

A. Positions of the 3' processing sites in histone H3.3B gene



B. Histone H3.3B 3'UTR sequence

```
3851  CGCCGCATCC GTGGCGAGCG TGCTTAAGGT GGATCAGCAG GAACGCCAGA
3901  TTCGATCACT GTCGTCCATT CCCGAACAGT AGACCATTCA CGTACCGTTT
3951  TAGGAGAGGT GCACGCCAAG GAGCTATACC AGCGACACAT TGATCCAATC
4001  TATCATTCAT ATTTACATGT ATACATTTAT CATCCTTTTG TTGTTCAATC
4051  GAATCGTCAG CATGACGTTG TCGATAAGAA AAACACCACA CACAAAACAA   P1
4101  CACACAACAT TACTGTAATG CTTAAAAATT ATTTTTATAAC TCTTATTATT  SITE
4151  ACTATTATCA ATTAGTTTTT AAGTTTTTTT TCTCGCCTGT GGCTCGTCGG
4201  TGTCACCGTG TATGTTGCGC CGTGCTCCTT CTTCTCCAAA TCCGGGTCCA
4251  CCTCTTCGTT TTGGAAGGAG GGAATACCAC CGGGCGGCCG TTGTCCGTGT
4301  CGCCCTGCAA CATCCGCCAC CAACCACTCC AACAACTGTC CAGCTGGCGG
4351  GCTTCGGTTT CCAGGTAATC CGTGATCTTT AAGGCGCCTG CCTCGCCTTT
4401  ATCGCTTCGA TCGAGGAGAG GTTCACCATC TGGCTGCCGG GATGTGCGGA
4451  GCCTGGAGAT CGGGGCCAAT CTAACGTCGG CTGGACGTGG CAAGGCGGAC
4501  AGACGGTTGC GGACCGCCTG GCGGTGATTC CCGGATCTCT GGATGGGCGG
4551  CGAGGGCATG GGCCTCTGGC GATGCCACTG ACGGCACGCG GCTTTTATGA
4601  AGATAATATC CATATATAAC ACCGAAGACC GAATCGAGTA CACTTAACTT
4651  TACACTTACG TATATGCATA TATATACATA AATACATATA TATATAAATA   P2
4701  TATATATATA TATAAATATA AATATAAACC TATGCGGAAG GAAAAGAACT  SITE
4751  AGATATTAAT ACTCTTTCTT CTGTTCACTT TTCGTTACTG CAAATCATAA
4801  TCAGAGCATT ATATATATAT ATATATATAG AAGTTATATA TATATATATA
4851  CACATAACGA AGGGATGGGC GTGACAGAGT TAGTTGATTA AATAAGATTT
4901  TAGATTTACA TGCATTTTAT ACGTTTAGGC GGACGATAGA AGATGATTCT
4951  TCTTCGGGAG ATGAAAAAAA GATATAAAGA TGATTCAGCA ATCAGATTAA
5001  CTTACACATC TATAGATTTG GAAGAGCATC GAATAAAAAT AACCAAGACA
5051  GAAATCACAT AACAATACAA AATAAAAGCA AAAGCGCCGA GCAGAGCATT
5101  GAATATTTCC CAATATTTAT ATGTAGTTTG AAACACATGC AAAATAATCA
5151  ATTGTAAACG CCCATAAAAC ATGCATATAT ATATATATAA ATATACATAT
5201  AGCTGTAAAT TGGTGTCGAT AACTAGTTTC ACATTGAACA TTTAAACAAA
5251  TCGGACTAAA TATGTGTCTA TTCGTATCGT CGTAGCTTAC AAGAACCAAT
5301  ATGAAAACAG AAAACAAAAA ACGAATCCCC CACAAAACAC AAAAACAATT
5351  TACAACAACA AATCAGTGAC AGCAACAGCA AAAATAACTG GAAACATCAG   P3
5401  GCAAAAAAAA AAAAGAAAAT AAAAAATAAA CGGTTTTTTC TATAGAAGAA  SITE
5451  ACCAAAAAAA CCCAACACCA ACGAATTGTA TTTTCTGCGA AAGCGAGACG
5501  GGAATAGTGT GTACGAGCGG GCGGGATAGG TGGCGATGGT ACAGTTGACT
                                                  HincII
```

D.mel.	H3 3B P3	AATAAA CGGTTTTTTCTATA
D.hydei	H3.3B	AATAAA CGGTTTTTTCTAGA
D.mel.	H3 3A	AATAAAATCGATTCATTTTTGTTTGTA
D.hydei	H3.3A	AATAAAA ATACATTTTTGTTTTCT
Adenovirus L3 (strong)		AATAAA GGCAAATGTTTTTATTTGTA
Adenovirus L1 (weak)		AATAAAAAACTCACCAAGGCCATGGCA

Fig 2 Alignment of the sequences immediately downsteam of the polyadenylation signals of the *D melanogaster* and *D hydei* histone H3 3A and H3 3B genes and the adenovirus poly(A) sites L1 and L3 (Prescott and Falck-Pedersen 1994)

The results of the in vitro polyadenylation assay are shown in Fig 3 Both constructs derived from the P1 site, CON 1 and CON 4 were inactive in the polyadenylation assay This result is in agreement with the absense of the poly(A) signal in the region upstream of the P1 site Similarly CON 2 derived from P2 site, showed low activity in this assay The sequences around the P2 site, although rich in UA-repeats, do not contain a conventional poly(A) signal Surprisingly, "precleaved" substrate from P2 (CON 5) was polyadenylated in a way comparable to the control, U1 70K protein pre-mRNA Both constructs derived from the P3 (CON 3 and 6), which contains a duplicate poly(A) signal were very efficient in polyadenylation

4 Discussion

The results of the in vitro polyadenylation assay confirm the general rule that polyadenyation efficiency is dependent on the presence of the AAUAAA signal The substrates, derived from the P3 site, containing two such signals, were very efficient in this assay An interesting feature of this site, which might contribute to its efficient utilization, is the presence of a conserved U-rich element immediately downstream of the poly(A) signal (Fig 2) The sequences between the poly(A) signal and poly(A) site are normally not conserved However these sequences can influence the "strength" of a particular polyadenylation signal, as was shown by the comparison of a weak and a strong poly(A) sites from the

---

Fig 1 Processing sites in histone H3 3B 3 UTR
A Schematic representation of the *D melanogaster* histone H3 3B gene structure Positions of the processing sites P1, P2 and P3 are indicated The horizontal arrow indicates the start of transcription Black boxes represent translated part of the exons and open boxes represent untranslated part of the exons
B Genomic sequence of the 3 part of the *D melanogaster* histone H3 3B gene The stop codon is shown in bold and marked with STOP Long arrows above the sequence indicate positions of the primers used for making constructs for the polyadenylation assay The HincII site at the end of CON 3 is indicated below the sequence The 3-terminal nucleotides of the precleaved substrates are underlined The polyadenylation signals are shown in italics and underlined and the U-rich sequences are shown in italics The first A-nucleotides of the poly(A) tail at the experimentally determined polyadenylation sites are shown in bold and indicated with asterisks above the sequence

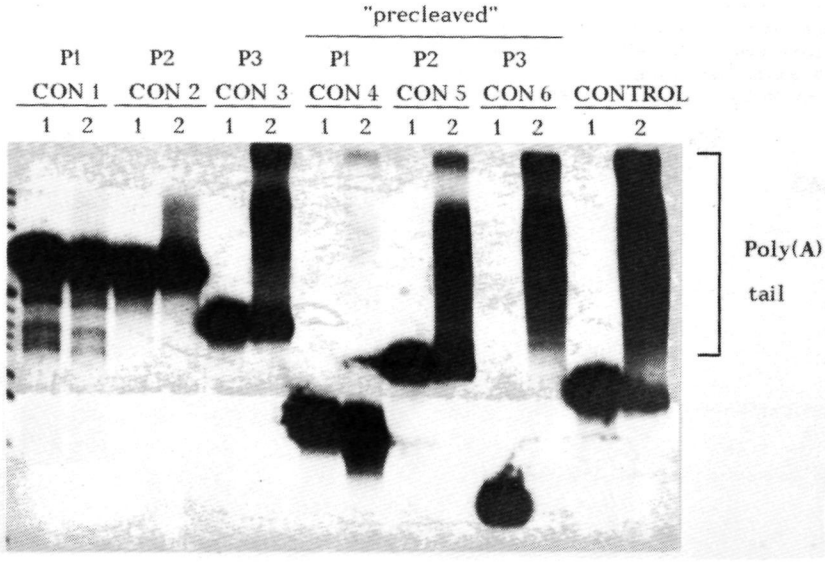


Fig. 3. In vitro polyadenylation of the histone H3.3B pre-mRNA substrates, corresponding to the three in vivo processing sites. For each substrate lane 1 is the input and lane 2 is the result of the incubation with the nuclear extract.

adenovirus major late transcription unit (Prescott and Falck-Pedersen, 1994). The sequence immediately downstream of the AAUAAA signal of the strong adenovirus poly(A) site, L3, also contains such a U-rich element, while a weak adenovirus poly(A) site, L1, does not (see Fig. 2).

P2 site contains several copies of a hexamer, AAUAUA (or UAUAAA), which is only by one nucleotide substition different from the consensus. Hexamers with these substitutions were shown to have very low activity as polyadenylation signals in the HeLa system (Wickens, 1990). However, the P2 site can apparently be recognized by the polyadenylation machinery, because it is quite efficient in polyadenylation assay, when a "precleaved" substrate is used. This might be due to the presence of several copies of the "mutated" hexamer. The activity of the P2 substrate, which is not "precleaved", is low, indicating that P2 is probably a weak processing site in vivo. It is likely to function only when the concentration of the processing factors is high. In such circumstances it may be favored compared to the strong P3 site, because it is localized upstream of it. This is apparently the case in embryos and in gonads, where P2 is preferentially utilized. This idea is supported by the fact that utilization of AAUAUA hexamer as the poly(A) signal was observed in the ovary-specific transcripts of the *exuperantia* gene (Hazelrigg and Tu, 1994) and testis-specific transcripts of the *radha* gene (Harhangi, unpublished).

The P1 site does not contain a conventional polyadenylation signal and it

is not processed by the in vitro polyadenylation system from HeLa cells. It is, however, efficiently used in *Drosophila* embryos and ovaries. It is possible that in *Drosophila* there are additional processing factors, which recognize it. Alternatively, the utilization of this site can be due to coupling between transcription and polyadenylation, which normally occurs in vivo, but does not happen in the in vitro polyadenylation system. It was shown that the utilization of a poly(A) site can be stimulated by downstream RNA polymerase II pausing and termination sites (Dorsett, 1990; Enriquez-Harris et al., 1991; Tantravahi et al., 1993). At present, no consensus for transcription termination sites is available and this idea cannot be substantiated by the analysis of the sequences downstream of P1. However, the fact that this processing site represents a special situation is supported by our observation that the poly(A)-addition sites are spread over a comparatively broad region (~50 nt, which is twice as much as was observed for the H3.3A genes and P2).

A good understanding of the regulation of the histone H3.3B pre-mRNA processing may only be achieved with the help of a *Drosophila* in vitro polyadenylation system. Also genetical tools, as they are available in *Drosophila*, might be quite useful for such an analysis. For example, it would be interesting to investigate if the utilization of the processing sites in the H3.3B gene is altered in the flies with mutations in the *supressor of forked* gene. This gene encodes a *Drosophila* homologue of the 77 kDa subunit of the mammalian cleavage stimulation factor, an essential component of the 3' processing machinery (Takagaki and Manley, 1994).

Acknowledgement
We are grateful to Koos Miedema for critical comments on this chapter.

Chapter IV

Two types of polyadenylated mRNAs are synthesized from *Drosophila* replication-dependent histone genes

Anna Akhmanova, Koos Miedema, Hannie Kremer[1] and Wolfgang Hennig

Department of Molecular and Developmental Genetics, Faculty of Sciences, Catholic University of Nijmegen, Toernooiveld, 6525 ED Nijmegen, The Netherlands

[1]Present address: Department of Human Genetics, Catholic University of Nijmegen, 6525 ED Nijmegen, The Netherlands

submitted for publication

Abstract

The polyadenylation of replication-dependent histone H2B, H3 and H4 mRNAs in *Drosophila melanogaster* was analysed. Two types of mRNAs, containing a poly(A) tail, can be detected in addition to non-polyadenylated messengers, which represent the majority of replication-dependent histone mRNAs. Firstly, conventional polyadenylation signals, localized downstream from the stem-loop region, are used to produce polyadenylated mRNAs. The messengers of this type, generated from the *D. melanogaster* H2B gene, are preferentially synthesized in the testis of the fly. Secondly, a novel type of polyadenylated histone mRNAs has been identified. These mRNAs, which are present in many different tissues and constitute a minor part of the total histone mRNA pool, contain a short poly(A) tail, added to the end of the 3' terminal stem-loop structure, which is in most cases lacking several nucleotides from its 3' end. The sites of polyadenylation within the stem-loop are not preceded by a normal polyadenylation signal. The possible functions of the polyadenylated histone transcripts are discussed.

1. Introduction

Replication-dependent histone genes in metazoans are in several ways different from the rest of the protein-coding genes. They do not contain introns and produce mRNAs without a poly(A) tail at the 3' terminus. Instead, these mRNAs end with a stem-loop structure, which has been highly conserved in evolution (Birnstiel et al., 1985; Marzluff, 1992). The expression of the replication-dependent histone genes is tightly linked to the S-phase of the cell-cycle, and its regulation occurs at multiple levels, such as transcription, mRNA processing and degradation (for review see Schümperli, 1988; Marzluff and Pandey, 1988; Osley, 1991). The 3' terminal stem-loop is essential for all postranscriptional steps of histone expression. It is necessary for efficient processing (Mowry et al., 1989; Vasserot et al., 1989), export of the mRNA into the cytoplasm (Eckner et al., 1991; Williams et al., 1994), loading onto polyribosomes (Sun et al., 1992) and regulation of mRNA stability (Pandey and Marzluff, 1987; Graves et al., 1987). In addition to the replication-dependent histone genes, there are also replacement, or basal histone genes, which are not dependent in their expression on the S-phase of the cell-cycle. They produce polyadenylated mRNAs and most of them, with the exception of the genes, incoding some linker histone variants and histone H2A.X, contain introns (for review see Zweidler, 1984; Schümperli, 1986).

Although the major portion of replication-dependent histone messengers are not polyadenylated, utilization of the poly(A) sites, localized downstream from the stem-loop sequence can result in formation of polyadenylated mRNAs. In some cases such mRNAs represent a very minor part of the total mRNA pool (Levine et al., 1987). In several histone H1 transcripts in chicken (Kirsh et al., 1989) and mouse (Cheng et al., 1989), as well in transcripts, encoding the histone H2A.X variant (Mannironi et al., 1989; Nagata et al., 1991) the two modes of the 3' processing occur with comparable efficiency in the same cell or tissue type. In all these cases the polyadenylated messengers behave as replacement

type histone mRNAs, which are not destabilized at the end of the S-phase, in contrast to the non-polyadenylated ones. The existence of such poly(A)-containing mRNAs is probably necessary to insure the basal synthesis of the respective histone variants (Kirsh et al., 1989; Cheng et al., 1989; Mannironi et al., 1989; Nagata et al. 1991).

Polyadenylated mRNAs, generated from replication-dependent histone genes, were also found in round spermatids of chicken (Challoner et al., 1989) and mouse (Moss et al., 1994). The addition of poly(A) tails to these messengers occurs downstream from the stem-loop region and the polyadenylation sites are not preceded by the conventional polyadenylation signal, AAUAAA. These polyadenylated mRNAs are translated in round spermatids, although the function of histone gene expression at this stage of spermiogenesis is unclear because histones are being replaced by transitional proteins and protamines.

While a wealth of information is available on the expression of histone genes in various vertebrates and sea urchins, comparatively little is known about the expression of these genes in *Drosophila*. However, due to the fact that *Drosophila* is becoming one of the most important systems in chromatin studies, detailed understanding of structure and function of the histone genes in this organism is essential. In the present work we have addressed the question of polyadenylation of *Drosophila* replication-dependent histone gene transcripts. We have found that the majority of these transcripts is not polyadenylated, as was shown before (Anderson and Lengyel, 1984), but that in addition two types of polyadenylated transcripts can be detected. Firstly, a small proportion of the histone mRNAs bears a short poly(A) tail which is added to the 3' terminus of a partially degraded stem-loop structure. Secondly, similar to what has been described in vertebrate systems, polyadenylation signals downstream of the stem-loop region can be utilised to generate mRNAs with a poly(A) tail. At least for the H2B gene this type of processing occurs preferentially in the testis.

2. Materials and Methods

2.1. Fly strains

The wild-type strains of *D. hydei* (Tübingen) and *D. melanogaster* Canton S were from our laboratory collection.

2.2. RNA isolation and Northern blotting

Total RNA was isolated by guanidinium-chloride method (Chirgwin et al., 1979) from whole male and female flies, from adult testes and ovaries and male and female carcasses (carcass is the remainder of the fly after the removal of the gonads. The poly(A)$^+$ RNA was prepared with the aid of a mRNA purification kit (Pharmacia). Approximately 20 µg of total or poly(A)$^-$ RNA or 1 µg of poly(A)$^+$ RNA per lane were separated on 2% agarose-formaldehyde gels and blotted to Hybond nylon membrane (Amersham). The blots were hybridized in 0.5 M phosphate buffer, pH 7.2, 7% SDS, 1% BSA and 1 mM EDTA at 60°C and washed in 0.04 M phosphate

buffer at the same temperature. Probes were labeled by PCR or by nick-translation in the presence of [α-^{32}P]dCTP. The position of the *D. melanogaster* probes, according to the numbering in Matsuo and Yamazaki, 1989, was as follows: histone H2B, positions 677-832; histone H3, 3303-3357; histone H4, positions 4316-4427 (see also Fig. 2). The *D. hydei* H2B probe was a subclone from the *D. hydei* histone repeat, positions 4357-4804 according to the numbering in Kremer and Hennig, 1990. The length of the transcripts was determined with the help of the GIBCO BRL RNA ladder and by comparing to the positions of the ribosomal RNAs.

2.3. Determination of polyadenylation sites by RT-PCR

Rapid amplification of the 3' ends of the cDNAs was performed essentially as described by Frohman, 1990. cDNA was synthesized from 10 μg of total RNA in the presence of 20 pmol of oligo(dT)-adapter primer with the sequence 5'-GCGGATCCGAATTCATCGT$_{17}$ and Moloney-Murine Leukemia Virus reverse transcriptase (GIBCO BRL). PCR was performed with a gene-specific primer and the oligo(dT)-adapter primer or an adapter primer with the sequence 5'-GCGGATCCGAATTCATCGT. The following gene-specific primers were used (the part of the primer, present in the corresponding gene, is underlined; see also Fig. 2): histone H2B, 5'-CATCTAGACCAAGTCGGGAGATCCAAAC; histone H3, 5'-GATGATCACGGCATTAACTTGC; histone H4, 5'-GATGATCACAGCCATGGATGTTGT; histone H3.3A, 5'-CAGGATCCGATAAGTAATCAATAC (positions 678-695 of the *D. melanogaster* histone H3.3A mRNA, Akhmanova et al., 1995). PCR products were separated on 2% or 3% agarose minigels and in some cases analysed by Southern blotting. The *D. melanogaster* histone H3.3A probe was as described by Akhmanova et al., 1995.

To compare non-polyadenylated and polyadenylated 3' ends, RNA ligase mediated rapid amplification of cDNA ends (RLM-RACE; Liu and Gorovsky, 1993) was used. Oligo 1 with the sequence 5'-TTTCTAGACTGGCCGTCGTTTTACA was ligated to the 3' ends of total RNA as described before by Liu and Gorovsky, 1993. The cDNA first strand was subsequently synthesized in the presence of oligo 2 with the sequence 5'-TGTAAAACGACGGCCAGT, which was complementary to the 3' part of the oligo 1. PCR was performed in the presence of a gene specific primer and oligo 2 or (for identification of polyadenylated molecules only) with the oligo 3 with the sequence 5'-TGTAAAACGACGGCCAGTCTAGAAATTT.

PCR products were subcloned in pBluescriptKS+ and sequenced either manually or with the aid of an Autoread kit (Pharmacia) on an ALFexpress automatic sequencer (Pharmacia).

3. Results

3.1. A minor part of the replication-dependent histone mRNAs contains a poly(A) tail

To assess the polyadenylation status of replication-dependent histone mRNAs from *D. melanogaster*, total and poly(A)$^+$ RNA from whole female and male flies was analysed by Northern blotting with probes, specific for histone H2B

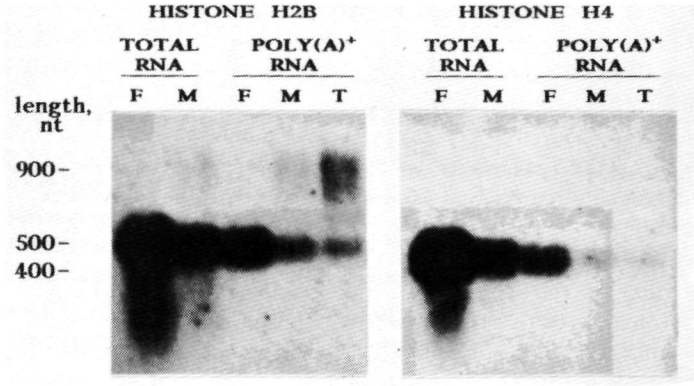


Fig. 1. Northern blot analysis of total RNA (20 μg per lane) and poly(A)$^+$ RNA (1 μg per lane) from *D. melanogaster* with *D. melanogaster* histone H2B and H4 probes. **F** - RNA from whole female flies, **M** - RNA from whole male flies, **T** - RNA from fly testes.

and H4 (see Fig. 1). Both probes contained a small carboxy-terminal part of the corresponding open reading frame and the 3' untranslated region (UTR) up to and including the stem-loop region. They displayed no cross-reaction with replacement-type mRNAs, such as the H4r mRNA (Akhmanova et al., 1996). Strong signals, corresponding to the 500 nucleotides (nt) histone H2B and 400 nt histone H4 mRNAs, were observed in total RNA samples. These transcripts are accumulated in the ovaries and stored in oocytes (Anderson and Lengyel, 1984). Therefore, they are more abundant in RNA from females than from males.

The transcripts of the same length could be also readily detected in poly(A)$^+$ RNA samples. Such samples, isolated by purification on oligo(dT)-cellulose are strongly enriched in polyadenylated mRNAs. However, they are still contaminated by non-polyadenylated RNAs, such as ribosomal RNAs. The latter is due to the high AT content of the *Drosophila* ribosomal RNAs. Such contaminants can amount to 50% of the poly(A)$^+$ fraction. Total RNA samples consist mostly of the ribosomal and transfer RNAs. The hybridization signals, obtained from 1 μg of poly(A)$^+$ RNA were several times weaker than those obtained from 20 μg of total RNA (4 to 10 times, depending on the sex of the flies and the histone gene analysed). If one assumes that polyadenylated mRNAs represent 1% of total RNA (0.2 μg per lane) and 50% of the poly(A)$^+$ RNA (0.5 μg per lane), the proportion of polyadenylated histone mRNA fraction may be estimated to be not more than 10% of the entire histone mRNA. This estimate is not accurate, since the retention of a particular mRNA molecule on the oligo(dT)-cellulose depends on several factors, such as the length of the poly(A) tail or occurence of internal A-rich stretches. However, it can be regarded as an indication that polyadenylated molecules represent a minor fraction of the short histone mRNAs.

Also with a histone H3 probe we could detect in the poly(A)$^+$ RNA frac-

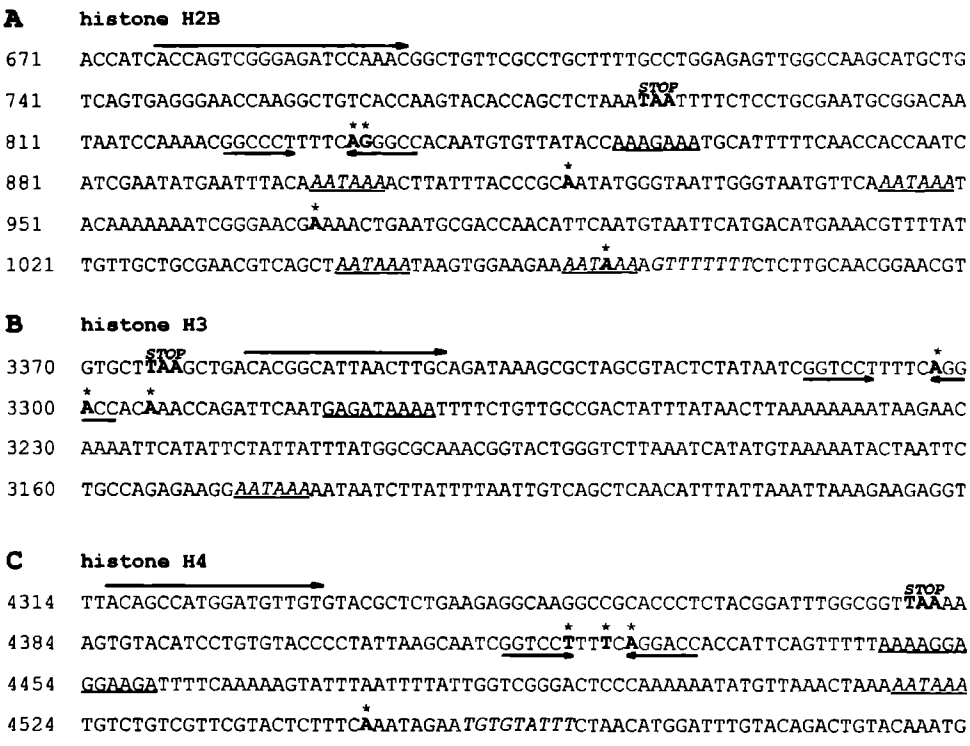


Fig. 2. The position of polyadenylation sites in the sequences of the 3'-terminal regions of the *D. melanogaster* histone H2B (**A**), H3 (**B**) and H4 (**C**) genes. The numbering corresponds to that in Matsuo and Yamazaki, 1989 and the coding strand of all three genes is shown. The stop codon is shown in bold and marked with *STOP*. Long arrows above the sequence indicate the positions of the gene-specific primers. The inverted repeats of the stem-loop regions are indicated by inverted arrows below the sequence and the purine-rich regions are underlined. The polyadenylation signals are shown in italics and underlined and the putative GU-rich elements are shown in italics. Experimentally determined polyadenylation sites are shown in bold and indicated with an asterisk above the sequence.

tion transcripts, similar in length to the major non-polyadenylated mRNA of this histone (not shown). Transcipts of the histone H1 and H2A genes were not analysed.

As the polyadenylated molecules were not much different in their length from the non-polyadenylated ones, we assumed that they beared short poly(A)-tails, added in the vicinity of the normal processing site, the stem-loop. This assumption was confirmed by a reverse transcription - polymerase chain reaction (RT-PCR) experiment with histone H2B, H3 or H4-specific primers (see Fig. 2) and oligo(dT). The major products of ~200 bp for histone H2B, ~100 bp for histone H3 and ~160 bp for histone H4 were obtained (see Fig. 3). The only exception were the longer (300-400 bp) products, obtained with the histone

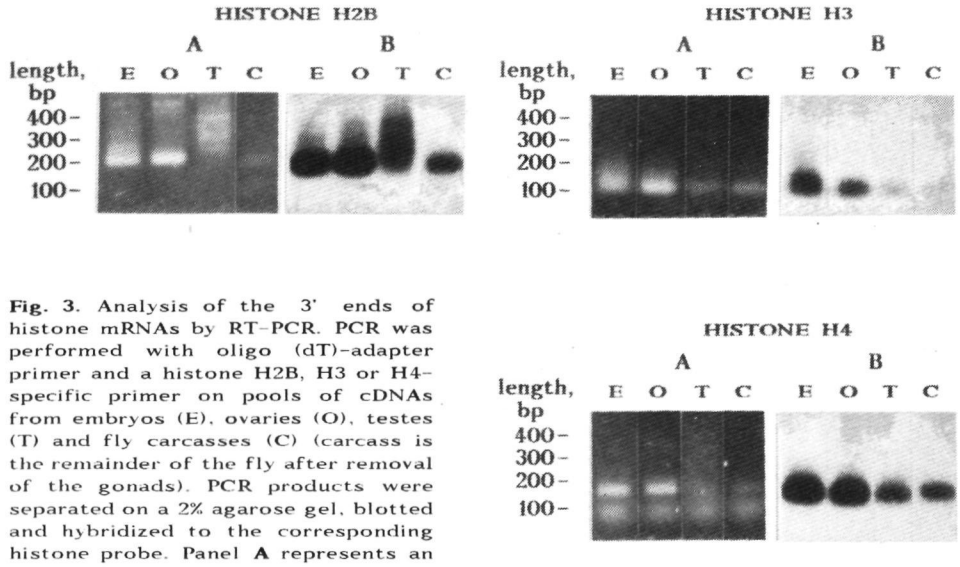


Fig. 3. Analysis of the 3' ends of histone mRNAs by RT-PCR. PCR was performed with oligo (dT)-adapter primer and a histone H2B, H3 or H4-specific primer on pools of cDNAs from embryos (E), ovaries (O), testes (T) and fly carcasses (C) (carcass is the remainder of the fly after removal of the gonads). PCR products were separated on a 2% agarose gel, blotted and hybridized to the corresponding histone probe. Panel **A** represents an ethidium bromide staining pattern of the gels, and panel **B** represents the Southern blots of the same gels.

H2B-specific primer from testis RNA, which are discussed below. The length of the PCR products corresponded to the distance between the gene-specific primer used and the stem-loop sequence of the particular histone gene, if the length of the oligo(dT)-adapter primer is taken into account. Such PCR products could not be obtained if only one primer (oligo(dT)-adapter primer or the gene specific primer) was used for PCR or if the amplification was performed with *D. melanogaster* genomic DNA as a template (data not shown). However, the products very similar to those shown in Fig. 3 were reproducibly obtained, using different cDNA pools, and also when the oligo(dT)-adapter primer in the PCR reaction was substituted for an oligonucleotide, which contains only the adapter part of the primer, but not the oligo(dT) part of it (see Materials and Methods). These controls indicate that the PCR products, obtained with oligo(dT), are unlikely to be PCR artifacts.

The PCR products were cloned and several subclones, corresponding to each of the three analysed histone genes were sequenced in order to determine the exact position of polyadenylation sites (see Fig. 4). As expected, the subclones contained the normal histone 3' UTR up to the stem-loop, and in two histone H3 subclones the poly(A) tail was present after the normal processing site. However, in most cases the subclones contained an incomplete stem-loop sequence. The 5' part of the stem was intact in all subclones analysed, while the poly(A) tail was added either to one of the loop nucleotides or in the middle of the 3' part of the stem. The preferred position of the poly(A) addition was after the 3' terminal loop nucleotide, C (or after the subsequent A residue, the two possibilities being impossible to distinguish in this type of experiment).

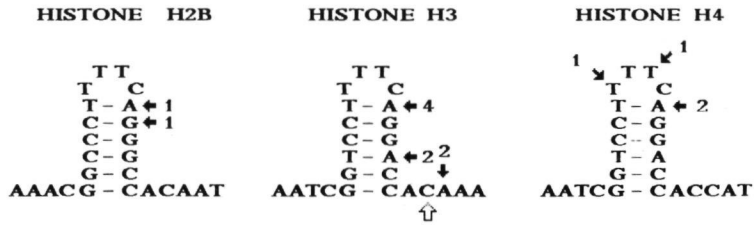


Fig. 4. Secondary structure of the 3'-terminal stem-loop regions of the *D. melanogaster* histone H2B, H3 and H4 genes is shown with polyadenylation sites indicated by thick black arrows. The number near each arrow is the number of independent subclones, with the poly(A) tail starting at that particular position. The open arrow below the histone H3 sequence indicates the experimentally determined 3'-terminal nucleotide of the non-polyadenylated histone H3 mRNA. Thin arrows indicate the positions corresponding to the ends of the decay products of the human histone H4 mRNA, described by Ross et al., 1986 (its stem-loop sequence is identical to that of the *D. melanogaster* histone H2B gene).

3.2. The distribution of the polyadenylation sites does not correspond to the distribution of the ends of the non-polyadenylated molecules

As the next step, we have analysed the 3' ends of the non-polyadenylated histone mRNAs, using the RNA ligase mediated rapid amplification of the cDNA ends (RLM-RACE) techique (Liu and Gorovsky, 1993). An oligonucleotide (oligo 1, see Materials and Methods) was ligated to the 3' ends of the RNA, and this RNA was used as a template for cDNA synthesis in the presence of the oligo 2, which was complementary to the 3' part of the oligo 1. The same oligo 2 and a histone H3 gene-specific primer were then used for PCR. Using this approach one can amplify the 3' ends of cDNAs irrespective of their polyadenylation status. The major product (see Fig. 5, lanes 1 and 3) corresponded to the 3' UTR sequence, containing an intact stem-loop structure. This was confirmed by sequencing of 8 independent subclones, which were all identical and were derived from mRNAs that were cleaved after the nucleotide indicated by an open arrow in Fig. 4 (the actual sequence is shown in Fig. 6). It is clear, therefore,

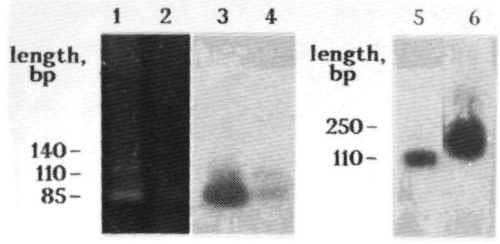


Fig. 5. Analysis of the length of nic RNA by ligating to it oligo 1 and reverse transcription in the presence of oligo 2, complementary to oligo 1 (see Materials and Methods). PCR was performed with the histone H3-specific primer and either oligo 2, which amplifies all cDNA ends (lanes 1 and 3) or oligo 3, which amplifies only polyadenylated cDNA ends (lane 2 and 4). The products were separated on a 3% agarose gel and hybridized to the histone H3 probe. An ethidium bromide staining pattern of the gel is shown in lanes 1 and 2 and the Southern blot of the same gel is shown in lanes 3 and 4. As a control the same pool of cDNAs was used for amplification with an H3.3A-specific primer and either oligo(dT) (lane 5) or oligo 2 (lane 6). The PCR products were separated on a 2% agarose gel and hybridized to the H3.3A probe (the results of the Southern blot are shown).

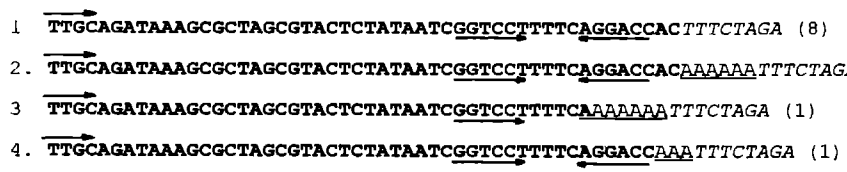


Figure 6 Sequences of the products obtained by RLM-RACE of the histone H3 transcripts derived from embryos PCR was performed with the H3-specific primer and oligo 2 (sequence 1) or with oligo 3 (sequences 2, 3 and 4) The part of the sequence present in the histone H3 gene, is shown in bold The horizontal arrows above the sequences indicate the end of the gene-specific primer, used for PCR Sequence of the 5 part of the oligo 1, used in the RNA ligation step is shown in italics The inverted repeats of the stem-loop regions are indicated by inverted arrows below the sequence The poly(A) tail sequences are underlined Number of independent subclones with a particular sequence is indicated in parentheses

that the distribution of the polyadenylation sites in the stem-loop sequence is not just a reflection of the 3' end constitution of the non-polyadenylated molecules

3.3 Poly(A) tails added to the stem-loop sequence are short

Comparison of the lengths of histone transcripts from total and poly(A)$^+$ RNA samples suggested that the poly(A) tails added to the hairpin structure, cannot be long (Fig 1, see above) This was confirmed by RLM-RACE As was mentioned before the use of a gene-specific primer and the oligo 2 allows to amplify ends of all cDNAs, whether polyadenylated or not However, when the oligo 2 is substituted for oligo 3, which is fully complementary to the oligo 1, used in the initial RNA ligation step, but contains 3 additional T residues at its 3' end only the cDNAs, which do contain at least a short poly(A) tail, can be amplified The PCR products, obtained with the oligo 3 and the histone H3-specific primer are quite similar to those, obtained with oligo 2 and histone H3-specific primer (Fig 5, lanes 2 and 4) Taking into account that most of the polyadenylated mRNAs are lacking a few nucleotides at the 3' end of the stem-loop one can conclude that the majority of the molecules have a very short tail (less than 10 nt) This was confirmed by cloning and sequencing several PCR products obtained by RI M-RACE In three such subclones analysed, poly(A) tails of 3 6 and 7 nucleotides were observed and in two of these three subclones the stem-loop sequence was partially truncated (Fig 6) These cDNA fragments obtained by PCR without involving oligo(dT), provide direct evidence for the existence of histone mRNAs with short poly(A) tails

The length of the longest PCR products obtained with oligo 3 (Fig 5, lane 4) exceeds the length of the major PCR product obtained with oligo 2, by not more than 25-30 bp This indicates that the poly(A) tails, added to the stem-loop sequence of histone mRNAs are not longer than 30 nt This is in contrast to the length of the poly(A) tails in the normal polyadenylated mRNAs, which is usually

in the range of 100–300 nt (Brawerman, 1981). We confirmed this by using the previously described procedure to amplify the 3' ends of the histone H3.3A mRNAs. Histone H3.3A gene encodes a replacement type histone and its mRNA is processed via the normal polyadenylation pathway (Akhmanova et al., 1995). The 5' end of the H3.3A-specific primer used was ~70 bp upstream from the polyadenylation site and the amplification with this primer and an oligo(dT)-adapter (35 nt long) results in products of ~110 bp (Fig. 5, lane 5). The PCR products obtained with H3.3A-specific primer and oligo 3 were in the range 120–250 bp (Fig. 5, lane 6). Taking into account the length of the ligated primer (25 nt), one can conclude that the poly(A) tails of the histone H3.3A mRNAs are 20–150 nt long.

3.4. Long polyadenylated transcripts of histone H2B are synthesized preferentially in the testis

In addition to short (~500 nt) histone H2B transcripts, longer (~600–900 nt) histone H2B mRNAs are present in RNA samples from male flies (Fig. 1). These mRNAs are polyadenylated and are derived from the testis, where they constitute the predominant part of the polyadenylated histone H2B transcripts (lane T in Fig. 1). An RT-PCR experiment with oligo(dT) and the H2B specific-primer on the testis RNA resulted in products, that were 100 to 200 bp longer, than those obtained from other tissues (Fig. 3). This indicates that the histone H2B mRNAs are polyadenylated in the testis at positions 100–200 nt downstream of the stem-loop.

The analysis of the 3' flanking sequence of the histone H2B gene revealed the presence of four polyadenylation signals, localized 60 to 230 bp downstream of the H2B stem-loop sequence (see Fig. 2A). The position of polyadenylation sites were determined by cloning and sequencing of the testis histone H2B RT-PCR products. As one can see from Fig. 2A, these polyadenylation sites are localized approximately 20 bp downstream from either one of the three polyadenylation signals. The positions of these sites are in good agreement with the length of the polyadenylated transcripts, observed on the Northern blot, if a poly(A) tail of ~150 nt long is taken into account. Small amount of such longer H2B transcripts could also be detected by RT-PCR in the RNA samples from ovaries and embryos (see Fig. 3), where they constitute a minor proportion of the pool of the histone H2B mRNAs.

There is also a single polyadenylation site in the 3' flanking region of histone H4 gene. It is utilized to generate a polyadenylated mRNA which is present in very low abundance, mainly in the testis. It could be detected by RT-PCR, but not by Northern blotting. Subcloning and sequencing of the testis RT-PCR product indicated the presence of a polyadenylation site ~30 bp downstream of the poly(A) signal (see Fig. 2C).

For the cell-cycle regulated histone H3 no long transcripts, resulting from the utilization of the downstream polyadenylation sites could be detected, although several such sites are present to the 3' of the histone H3 stem-loop region (the most upstream polyadenylation site is indicated in Fig. 2B).

3.5. Polyadenylation signals downstream of the replication-dependent histone H2B gene are not conserved in *Drosophila* evolution

The sequences of the replication-dependent histone genes are available from another *Drosophila* species, *D. hydei* (Kremer and Hennig, 1990), which is separated from *D. melanogaster* by 60 million years of evolution (Throckmorton, 1975). Contrary to the situation in *D. melanogaster* no polyadenylation signals are found downstream of the histone H2B gene of *D. hydei*. We analysed the transcripts of the H2B gene of *D. hydei* RNA by Northern blotting (see Fig. 7). The major type of H2B mRNAs both in testes and in carcasses is 500 nt long. Similar as in *D. melanogaster*, short histone H2B transcripts are also present in the poly(A)$^+$ RNA fraction. Such polyadenylated mRNAs are more abundant in the RNA from testis than from male fly carcasses. Additional weak signals corresponding to RNA species of 700 nt are present in both testis and carcass poly(A)$^+$ RNA samples. Finally, a testis-specific transcript of 1200 nt and a carcass-specific transcript of 1500 nt, were observed. These RNAs are unlikely to be synthesized from the replication-dependent histone H2B gene, because the strength of the signals is reduced with increasing hybridization stringency. This indicates that they have low homology to the probe and are probably derived from replacement-type H2B genes. Also by RT-PCR we could detect no long polyadenylated messengers of the replication-dependent histone H2B gene in *D. hydei* RNA samples. This is in agreement with the absence of polyadenylation signals in the 3' flanking sequences of the H2B gene in this species.

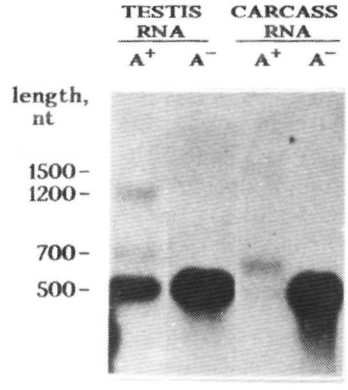


Figure 7. Northern blot analysis of testis and male carcass poly(A)$^+$ RNA (1 µg per lane) and poly(A)$^-$ RNA (20 µg per lane) from *D. hydei* with *D. hydei* histone H2B probe.

4. Discussion

4.1. Polyadenylation of the 3' terminal stem-loop sequence of histone mRNAs

In this study we report that a part of replication-dependent histone messengers in *Drosophila* bears a poly(A) tail added to one of few positions in their 3' stem-loop structure. Such polyadenylated molecules constitute a minor part of the pool of the histone mRNAs and their proportion can vary to a certain extent in different tissues.

The poly(A) tails added to the 3' terminus or to one of the internal positions of the stem-loop structure of the *Drosophila* replication-dependent histone

mRNAs are clearly distinct from the normal poly(A) tails of other mRNAs. These tails are quite short (up to 30 nt) and they are not preceded by a conventional polyadenylation signal. The positions of poly(A) sites in the stem-loops of histone H2B, H3 and H4 mRNAs are similar, being confined to the loop and the 3' half of the stem. Since the upstream 3' UTR sequences, with the exception of the stem-loop itself and a few nucleotides around it, are different, it is reasonable to assume that the position of the poly(A) addition is somehow directed either by sequence or the secondary structure of the stem-loop or by proteins bound to it. It seems not unlikely that the poly(A) tail addition occurs after the normal histone mRNA 3' processing, either in the nucleus or in the cytoplasm.

Comparison of the distribution of the polyadenylation sites to the distribution of the ends of the non-polyadenylated molecules showed that our data cannot be interpreted as a simple random addition of the poly(A) tails to the minor proportion of all the histone mRNAs. While the non-polyadenylated molecules contain an intact stem-loop sequence, the latter is partially absent in most of the polyadenylated ones.

One possible explanation of our data is that an addition of the short poly(A) tails occurs concomitantly with histone mRNA degradation and that the relatively stable degradation products, which may serve as a good substrate for polyadenylation, receive a poly(A) tail. It is noteworthy in this respect that the preferred position of the poly(A) tail, the CA dinucleotide at the base of the loop, is the dinucleotide which is found in the majority of the normal polyadenylation sites (Sheets et al., 1990). However, the localization of the polyadenylation sites within the stem-loop does not support this hypothesis, being in no apparent correlation with positions of the termini of the products of the initial steps of histone mRNA degradation. As was demonstrated for the human histone H4 mRNA, the degradation proceeds from the 3' terminus and discrete decay products, lacking 5 and 12 nucleotides from the 3' end can be observed (Ross et al., 1986; see thin arrows in Fig. 4). We found no polyadenylated mRNAs, corresponding to such decay products. One should bear in mind, however, that the degradation of *Drosophila* histone mRNAs might proceed in a way, different from that of mammals or that the -5 and -12 degradation products are especially poor substrates for polyadenylation.

Short poly(A) tails (less than 20 nt), added to the 3' end of an intact stem-loop sequence, were observed before for histone mRNAs in *Xenopus* oocytes (Ballantine and Woodland, 1985). These tails were removed upon fertilization and they were implied in the control of histone mRNA translation (Ruderman et al., 1979). The polyadenylated histone mRNAs that we observed in *Drosophila* are quite distinct, however, because in most cases at least a part of the stem-loop is missing. The normal stem-loop structure is essential for different aspects of histone mRNA metabolism and is apparently bound to protein factors, which mediate these functions. One such factor, a 45 kDa stem-loop binding protein, has been identified in mammalian cells and the sequence requi-

rements for its binding to the histone mRNA 3' end have been determined (Williams and Marzluff, 1995). An intact stem and sequences 5' and 3' of it are essential for its binding. It is reasonable to assume that at least a part of polyadenylated molecules will not bind this protein any more and will therefore be excluded from the normal pathway of histone mRNA turnover. On the other hand, presence of poly(A) could stabilize these mRNAs or stimulate their translation, as it was shown for polyadenylated mRNAs in a variety of systems (for reviews, see Munroe and Jacobson, 1990; Jackson and Standart, 1990; Sachs, 1993).

4.2. Utilization of the downstream polyadenylation signals

Similar to what has been observed in vertebrate systems, poly(A) signals localized 3' to the stem-loop region can be utilized to generate polyadenylated messengers in *Drosophila*. This utilization is not very efficient, probably because of the absense of introns in replication-dependent histone genes. It has been suggested that splicing and polyadenylation are coupled processing events (see Nesic and Maquat, 1994 and references therein). It was shown that introducing an intron into a gene containing the histone 3' end interferes with histone 3' end formation and favors usage of cryptic downstream polyadenylation sites (Pandey et al., 1990).

Comparing the amount of polyadenylated mRNAs synthesized from the histone H2B and H4 genes, one can conclude that polyadenylation efficiency correlates with the number of poly(A) signals. Only one such signal is present downstream of the H4 gene and the polyadenylated transcripts resulting from its utilization are detectable only by RT-PCR. In contrast, considerable amount of polyadenylated messengers are generated from the H2B gene, which contains four poly(A) signal downstream of the stem-loop. At least three of these four signals are used. Their utilization displays no apparent correlation with the presence of downstream GU or U-rich sequences, which have been shown to stimulate 3' processing (Wahle and Keller, 1992; Wahle, 1995; Keller, 1995). For instance, there is a GU_7 sequence 25 bp to the 3' of the third polyadenylation signal in H2B gene (see Fig. 2A), while no GU or U-rich elements can be found downstream of the other three polyadenylation signals. However, there is no indication, either from the Northern blot or from the RT-PCR, that the third polyadenylation signal is preferred to the other three (see Fig. 1 and 2A).

Remarkably, histone H2B mRNA polyadenylation occurs preferentially in the testis. Also in chicken (Challoner et al., 1989) and mouse (Moss et al., 1994) the synthesis of polyadenylated mRNAs from replication-dependent histone genes is observed in the male germ line cells. It seems, therefore, that formation of such type of histone transcripts represents a general feature of spermatogenesis both in vertebrates and invertebrates. However, histone mRNA polyadenylation in *Drosophila* is in several aspects different from that observed in chicken and mouse. Firstly, in the two latter systems the poly(A)-containing histone mRNAs were specifically expressed postmeiotically, i.e. in round spermatids (Challoner et al., 1989; Moss et al., 1994). In *Drosophila* no transcription

occurs after meiosis (Hennig, 1967), and the polyadenylated H2B mRNAs must be synthesized in spermatogonia or spermatocytes. It is possible, of course, that these transcripts are stored to be translated after meiosis. Secondly, both in chicken and mouse the polyadenylation sites of histone mRNAs are not preceded by a normal polyadenylation signal (Challoner et al., 1989; Moss et al., 1994). The authors suggested that cryptic signals, deviating from the consensus might be used. In *Drosophila*, however, a strict correlation was observed between polyadenylation and the presence of normal polyadenylation signals.

The function of these polyadenylated transcripts in the male germ cells is not clear. It is possible that they are necessary for expression of histone proteins after the cessation of DNA replication in the course of spermatogenesis. As it has been mentioned above, due to the presence of a poly(A) tail such mRNAs are usually more stable in the absence of DNA synthesis, than normal non-polyadenylated histone messengers. This interpretation is contradicted by the apparent lack of conservation of downstream polyadenylation signals in replication-dependent histone H2B genes of different *Drosophila* species.

On the other hand, the synthesis of such polyadenylated mRNAs might be spurious and result from the special properties of the polyadenylation machinery in the male germ cells, such as a less tight coupling of polyadenylation and splicing. It is known that the effect of transcriptional regulation on the level of histone gene expression in the different phases of cell-cycle are quite moderate and that the main regulatory events occur at the level of 3' processing and mRNA stability (Schümperli, 1988; Marzluff and Pandey, 1988; Osley, 1991). It is reasonable to assume, therefore, that also during spermatogenesis no efficient transcriptional silencing of replication-dependent histone genes takes place. Under these circumstances an increase of polyadenylation efficiency of histone mRNAs will result in a higher level of such mRNAs in male germ cells compared to somatic tissues. Such an interpretation is supported by the fact that in mouse spermatids poly(A)$^-$ transcripts are synthesized from the histone H2A gene and translated along with the polyadenylated ones (Moss et al., 1994). Moreover, in this system the non-polyadenylated mRNAs of all four core histones are more prominent than polyadenylated ones, which indicates that the latter play a minor role in histone gene expression.

In contrast to the histone H2B and H4 genes, we could observe no utilization of the downstream poly(A) signals of the histone H3 gene. Our preliminary data suggest that at least in the testis this might be accounted for by the specific silencing of this gene in primary spermatocytes.

Acknowledgements. We are very grateful to Denise Vlooswijk for the help with analysing the PCR product subclones. We thank Petra Bindels, Harry Harhangi and Wenli Gu for help and valuable discussions. This research was supported by the European Community Human Capital Programme (Networks), contract No. CHRX-CT94-0511.

Chapter V

Naturally occurring testis-specific histone H3 antisense transcripts in *Drosophila*

Anna Akhmanova, Hannie Kremer[1], Koos Miedema and Wolfgang Hennig

Department of Molecular and Developmental Genetics, Faculty of Sciences, Catholic University of Nijmegen, Toernooiveld, 6525 ED Nijmegen, The Netherlands

[1] Present address: Department of Human Genetics, Catholic University of Nijmegen, 6525 ED Nijmegen, The Netherlands

submitted for publication

Abstract

While analysing the transcription of the cluster of cell-cycle regulated histone genes in *Drosophila hydei* we have found transcripts, that were spanning both histone H3 and H4 genes and were antisense for histone H3. As the two histone genes are in opposite orientation, these transcripts contained the sense strand of the histone H4 gene. Such transcripts were present in both poly(A)$^+$ and poly (A)$^-$ RNA fractions. The polyadenylated molecules contained a poly(A) tail at the 3' end of the stem-loop structure, which is characteristic for cell-cycle regulated histone mRNAs. The antisense RNA of histone H3 is synthesized exclusively in testes and is restricted to the nuclei of primary spermatocytes. Possible functions of this RNA are discussed.

1. Introduction

Histones are small proteins that bind to the DNA of all eukaryotic cells and form nucleosomes, the basic structural units of chromatin. In metazoa, histones are encoded by several types of genes (Zweidler, 1984; Schümperli, 1986). The replication-dependent histone genes are expressed during the S phase of the cell cycle and are responsible for the major part of the histone synthesis in dividing cells. These genes produce predominantly short non-polyadenylated transcripts, which have a characteristic hairpin structure at their 3' ends (Birnstiel et al., 1985). The regulation of replication-dependent histone gene expression was shown to occur at multiple levels, such as transcription, 3' processing and mRNA stability (Schümperli, 1988; Marzluff and Pandey, 1988). Histone genes of another type, replacement genes, are not linked in their expression to the S phase and produce polyadenylated messengers (Zweidler, 1984; Schümperli, 1986). Their function and mode of regulation are still poorly understood. Finally, while most of the histone genes are expressed in cells of all types, also certain tissue-specific histone variants are found (Schümperli, 1986).

The histone expression patterns are especially complex during male germ line development. During spermatogenesis of most animal species chromatin organization undergoes drastic rearrangements (Poccia, 1986; Hecht, 1989a). Somatic histones are substituted by testis-specific histones and/or by small highly basic proteins such as protamines (Hecht, 1989b). It is clear, therefore, that as far as histone expression is concerned, the male germ line represents a special situation and is likely to possess special regulatory mechanisms.

Although *Drosophila* is one of the best studied model systems in chromatin research, surprisingly little is known about chromatin reorganisation during spermatogenesis. Visualisation of DNA by fluorescent microscopy revealed several cycles of chromatin condensation and decondensation in male germ line cells of *D. hydei* (Kremer et al., 1986). Histone H1 is apparently eliminated from the chromatin early in spermatocyte development, while histone H2A persists in the spermatid nuclei until late elongation (Kremer et al., 1986). Cytochemical experiments with *D. melanogaster* implicated that lysine-rich nuclear proteins are substituted for arginine-rich ones at late postmeiotic stages (Hauschteck-Jungen

and Hartl, 1982). However, no *Drosophila* testis-specific histones were described so far. Among the replacement variant genes, histone H3.3 was shown to be strongly expressed in the testis, at least at the mRNA level (Akhmanova et al., 1995).

We have undertaken a systematic analysis of histone transcription in testes. *D. hydei*, rather than *D. melanogaster*, was our model of choice, because of the larger size of the testis and the ordered arrangement of male germ cells in this species. We have identified a novel testis-specific antisense transcript of histone H3, which displays an unusual pattern of polyadenylation. With an improved protocol for in situ hybridization of oligonucleotide probes to *Drosophila* testis we have shown that this transript is localized in the spermatocyte nuclei. This transcript is likely to have a regulatory function.

2. Materials and Methods

2.1. Fly strains

The wild-type strains of *D. hydei* (Tübingen) and *D. melanogaster* Canton S were from our laboratory collection.

2.2. Isolation of RNA and Northern blotting

Total RNA was isolated from different samples by the guanidinium-chloride method (Chirgwin et al., 1979). For separation of the poly(A)$^+$ and poly(A)$^-$ RNA, a mRNA purification kit (Pharmacia) was used according to the protocol of the manufacturer. Approximately 20 μg of poly(A)$^-$ RNA or 1 μg of poly(A)$^+$ RNA per lane were electrophoresed on 2% agarose-formaldehyde gels and transferred to Hybond nylon membrane (Amersham). The blots were hybridized in 0.5 M phosphate buffer, pH 7.2, 7% sodium dodecyl sulfate (SDS), 1% bovine serum albumin (BSA) and 1 mM EDTA at 60°C. Non-stringent washing was done in 0.3 M phosphate buffer, pH 7.2, 1% SDS at 50°C and stringent washing in 0.02 M phosphate buffer at 70°C. Double-stranded DNA probes were labeled by nick-translation in the presence of α-[^{32}P]dCTP. The position of the probes (according to the numbering in Kremer and Hennig (1990)) was as follows: H3 probe, positions 2160-2581, H4 probe, positions 2648-3201. Strand-specific histone H3 probes were generated by in vitro transcription with the T7 RNA polymerase in the presence of α-[^{32}P]UTP of a linearized plasmid, containing a fragment of the H3 gene, corresponding to the nucleotide positions 2252-2534.

2.3. cDNA isolation by RT-PCR

The RT-PCR reactions were performed essentially as described earlier (Akhmanova et al., 1995). The 3' part of the *D. hydei* H3 antisense cDNA was amplified with primers H3-II ((Akhmanova et al., 1995), see Fig. 2A) or HAS4 with the sequence 5'-GTGCGATTTGTATATGTACAAGGCAGA and oligo(dT)-adapter primer on the pool of testis cDNA, synthesized in the presence of oligo(dT). For the amplification of the 5' end of this cDNA, RT reaction on *D. hydei* testis RNA was primed with the oligonucleotide H3-I (Akhmanova et al., 1995), then oligo(dC)

tails were added, and the tailed cDNAs were amplified with oligo(dG)-adapter primer and the primer H3-IV (Akhmanova et al., 1995). PCR products were cloned and sequenced.

2.4. In situ hybridization to testis squashes

Oligonucleotides were labeled by tailing with terminal transferase and digoxigenin-labeled dUTP (Boehringer Mannheim) in the presence of dATP according to the protocol of the manufacturer. The following oligonucleotides were used (see Fig. 2A): H3 antisense probe: HAS1 5'-ATAAGGCTTGCCTCCAGTTGATTTGCGAGC, HAS2 5'-CGACGAATTTCACGCAGAGCAACAGTGCCA, HAS3 5'-CCAGTTGAATGTCCTTGGCGCATGATTGTGA and HAS4 (see above). H3 sense probe: HS1 5'-GCTCGCAAATCAACTGGAGGCAAGGCTC, HS2 5'-TGGCACTGTTGCTCTGCGTGAAATTCGTCG, HS3 5'-TCACAATCATGCCCAAGGACATTCAACTGG and HS4 5'-TCTGCCTTGTACATATACAAATCGCAC. To detect H3.3A mRNA we used separately a 5' untranslated region (UTR) oligonucleotide with the sequence 5'-CAATTAGAGCACTTTTGTCACGGATAAAGTGACGTCCG and a 3' UTR oligonucleotide with the sequence 5'-TCTTATGAAACTTAAGGCTAAGCAGACACAGCAGC.

Testes of *D. hydei* were prepared in testis buffer (Hennig, 1967), squashed gently under a siliconized coverslip and heat-fixed for 3 minutes at 50°C in a humid chamber. The squashes were frozen in liquid nitrogen, the coverslip was removed and the slides were transferred immediately into chloroform for 3 minutes. The squashes were subsequently fixed in 4% paraformaldehyde/phosphate-buffered saline (PBS; contains 137 mM NaCl, 27 mM KCl, 1.5 mM KH_2PO_4 and 6.5 mM Na_2HPO_4 (pH 7.4)) for 10 minutes and washed once with PBS for 5 minutes and twice with 2xSSC for 5 minutes (1xSSC is 150 mM NaCl, 15 mM sodium citrate, pH 7.2). Prehybridization was performed at 37°C for 30-60 minutes in hybridization buffer (4xSSC, 10% dextran sulfate, 0.02% Ficoll 400, 0.02% polyvinyl pyrolidone, 0.02% BSA, 2 mM EDTA, 50% formamide). 40 µl of hybridization solution (hybridization buffer containing 100-300 pmol/ml of a digoxigenin-labeled oligonucletide mix) was applied to each squash and the hybridization was carried out overnight in a humid chamber at 37°C. Slides were washed once in 2xSSC at room temperature for 3 minutes, in 0.2xSSC at 37°C two times for 15 minutes and then in PBT buffer (2 mg/ml of BSA and 0.1% Triton X-100 in PBS) for 30 minutes at room temperature. Squashes were incubated for 90 minutes in 1:50 dilution of polyclonal sheep anti-digoxigenin-Fab fragments conjugated to alkaline phosphatase (Boehringer Mannheim) in PBT at room temperature. Slides were washed twice with PBT for 10 minutes, once with PBS for 10 minutes and then equilibrated for 10 minutes in the NBT buffer (100 mM Tris-HCl pH 9.5, 100 mM NaCl, 50 mM $MgCl_2$). Staining was performed in a solution, containing 5-bromo-4-chloro-3-indolyl phosphate 4-toluidine salt at 1 mg/ml and 4-nitroblue tetrazolium chloride at 2 mg/ml in NBT-buffer for 5-8 hours at room temperature, and the reaction was stopped in PBS.

Photographs were taken with a Zeiss photomocroscope III on Kodak Ektachrome 400 film.

3. Results and discussion

3.1. Identification of histone H3 antisense transcripts

In order to characterize the transcription of the histones H3 and H4 in *Drosophila* testis, we have probed Northern blots with poly(A)$^+$ and poly(A)$^-$ RNA from testes and carcasses (remainder of the fly after removing the testis) with double-stranded DNA fragments, containing the *D. hydei* cell-cycle regulated histone H3 and H4 genes (Fig. 1A and B). As was expected for cell-cycle regulated histones, both probes hybridized to short non-polyadenylated transcripts of approximately 500 nucleotides (nt) for histone H3 and of 400 nt for histone H4. In addition, the histone H3 probe cross-reacted with histone H3.3A and *H3.3B* mRNAs, which are very abundant in the poly(A)$^+$ samples. The identity of these transcripts was established before with the help of 3' untranslated region (UTR) probes (Akhmanova et al., 1995).

The histone H4 probe also recognized a polyadenylated transcript, generated from a histone replacement gene, *H4r* (Akhmanova et al., 1996). All these transcripts displayed no testis specificity, as they were present in the RNA isolated from fly carcasses as well.

In addition, both histone probes hybridized to a testis-specific RNA species of 1300 nt, present in the poly(A)$^+$ as well as in the poly(A)$^-$ RNA fractions.

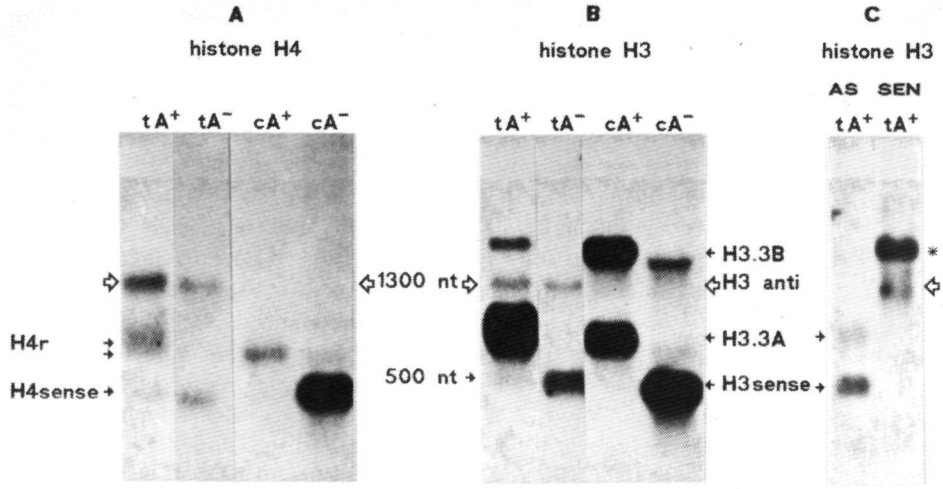


Fig. 1. Northern blots with testis (**t**) and carcass (**c**) RNA. About 1 μg of poly(A)$^+$ and 20 μg poly(A)$^-$ RNA per lane were loaded. **A.** Blots were hybridized with the double strand histone H4 DNA probe and washed non-stringently. **B.** Blots were hybridized with the double strand histone H3 DNA probe and washed non-stringently. **C.** Blots were hybridized with the antisense (**AS**) and sense (SEN) histone H3 RNA probes and washed stringently. Open arrows indicate the position of the histone H3 antisense transcripts. The signal, resulting from the cross-reaction of the sense RNA probe to an RNA species of ~2000 nt is indicated by an asterisk.

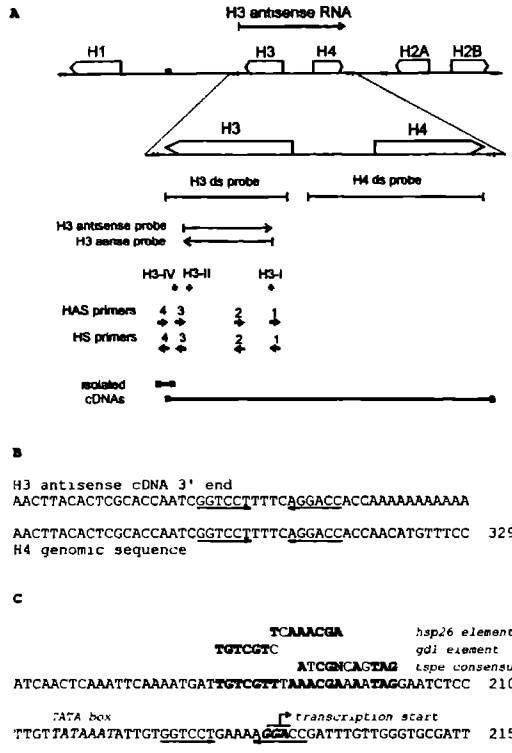


Fig. 2. A. Organization of the *D. hydei* cell-cycle regulated histone genes within a repeat unit and the position of the histone H3 antisense transcript. Histone open reading frames are indicated by open arrows and the positions of the inverted repeats by small inverted arrowheads. The TTGTCAAAT element is represented by a black box. Relative positions of double-strand (ds) and single strand (sense and antisense) probes and primers, used in this study, are indicated. Positions of the cDNAs for the antisense transcript, isolated by RT-PCR, are shown. **B.** Sequence of the 3' end of the histone H3 antisense cDNA and the corresponding genomic sequence. Numbering of the nucleotides corresponds to that in Kremer and Hennig (1990). Position of the 3' terminal inverted repeat of histone H4 gene is indicated by horizontal arrows. **C.** Sequence around the transcription start site of the histone H3 antisense transcript. Numbering of the nucleotides corresponds to that in Kremer and Hennig (1990). Transcription start is shown in bold and italic. Similarities to the known testis promoter elements, including testis-specific promoter element *(tspe)* consensus (Lankenau et al. 1994) are shown in bold. Sequences, complementary to the 3' inverted repeat of the histone H3 gene are indicated by horizontal arrows.

Because of this unusual property we expected both probes to hybridize to the same RNA molecules. The tandem genomic organization of cell-cycle regulated histone genes in *D. hydei* would permit such a transcription pattern (see Fig. 2A). However, as H3 and H4 mRNAs are transcribed from opposite DNA strands, a transcript spanning both genes would necessarily be antisense for one of them. We checked this possibility with strand specific histone H3 and H4 RNA probes. While no signals were obtained with the H4 sense probe, it was the sense, rather than the antisense histone H3 probe that hybridized to the testis-specific 1300 nt transript (see Fig. 1C). In addition, this probe also reacted with an RNA species of ~2000 nt (indicated by an asterisk in Fig. 1C). This additional signal was strongly reduced, but not eliminated completely by increasing the washing stringency. Such cross-hybridization is not uncommon for RNA probes, generated by in vitro transcription (Witkiewicz et al., 1993) and is probably caused by high stability of RNA-RNA hybrids. This signal was not due to the vector sequences, present in the probe, because those were identical for both sense and antisense H3 probes. It was never observed when the same fragment was used as a labeled DNA probe, rather than an RNA probe.

The results of Northern blotting indicated that the testis-specific transcript contained the antisense strand of histone H3 and the sense strand of histone H4 gene. This was confirmed by isolating this transcript from the pool of oligo(dT)-primed testis cDNA by using reverse transcriptase - polymerase chain reaction (RT-PCR). By using primers derived from the antisense strand of histone H3 (see Fig. 2A) in combination with oligo(dT) as the second primer we isolated cDNA fragments which contained the full length histone H3 open reading frame (ORF), the H3-H4 intergenic spacer, the full length histone H4 ORF and 3' UTR including the inverted repeat as well as a poly(A) tail added directly to the expected 3' processing site of histone H4 (Fig. 2A and 2B). These fragments could not be the result of amplification of genomic DNA contamination in the testis RNA, because the genomic DNA sequence immediately downstream of the histone H4 3'-terminal inverted repeat is not A-rich (see Fig. 2B).

The poly(A)-tails, present in the H3 antisense transcripts, can not be long, because the antisense RNAs, found in the poly(A)$^+$ and poly(A)$^-$ fractions have almost the same size (Fig. 1). Comparing the amounts of poly(A)$^+$ and poly(A)$^-$ RNA analysed and taking into account the fact that most of the poly(A)$^-$ fraction is represented by rRNA and tRNAs, we can conclude, that not more than 10% of all antisense transcripts are polyadenylated. We can not exclude, however, that all the antisense transcripts contain a poly(A) tail at the 3' end, but it is too short to ensure efficient binding to the oligo(dT) cellulose.

The fact that most of the histone H3 antisense RNA is apparently non-polyadenylated together with the structure of the 3' end of the polyadenylated molecules suggest that the 3' processing of these molecules occurs via the usual U7 snRNP-mediated pathway, as it is common for the cell-cycle regulated histone mRNAs (Mowry and Steitz, 1988; Marzluff, 1992). The addition of a poly(A) tail to these molecules is probably a secondary event, distinct from normal mRNA polyadenylation, since no polyadenylation signals are present in the 3' part of the transcript. This type of polyadenylation is unlikely to be unique for the H3 antisense transcripts, because some of the sense H3 and H4 transcripts are present in the poly(A)$^+$ fraction too (this is particularly apparent in Fig. 1C). The only other case where a short poly (A) tail, added to the histone 3' terminal stem-loop was found, is in *Xenopus* oocytes. In this system, the poly(A) is added during oogenesis and removed during oocyte maturation (Ballantine and Woodland, 1985). It was suggested to be involved in the regulation of translation (Ballantine and Woodland, 1985).

No significant polyadenylation of cell-cycle regulated histone mRNAs was found either in *Drosophila* oocytes or embryos (Anderson and Lengyel, 1984). From our Northern blots it is apparent that the extent of such polyadenylation is higher in testis than in carcass RNA.

3.2. Analysis of the promoter sequences of the antisense transcript

The 5' part of the cDNA was isolated by rapid amplification of the cDNA ends technique (Frohman, 1990). Surprisingly, the antisense transcription apparently

initiates within the 3'-inverted repeat sequence of the histone H3 (see Fig. 2C). Because we used C-tailing of the cDNA first strand it is impossible to say whether the first nucleotide is one of the two G residues (positions 2128 or 2129) or the subsequent A residue (Fig. 2C). This result is unlikely to represent an RT-PCR artifact, caused by the stability of the stem-loop structure in the antisense transcript, because even if the progression of the reverse transcriptase was hindered by the stem, the enzyme would stop at its 3' end rather than in middle of it.

Sequence analysis showed that there is a TATA-box at a position around -25 of the transcription start (Fig. 2C). We searched the upstream sequences for promoter elements identified in other genes, expressed in *Drosophila* testis. One sequence motif that acts as a promoter element to drive testis-specific transcription was originally identified in the β2-tubulin gene (Michiels et al., 1991). Similar sequence elements were subsequently found in a number of other testis-specific promoter regions, including that of the antisense transcript of the *D. hydei micropia* retrotransposon (Lankenau et al., 1994). A consensus sequence was proposed (Fig. 2C; Lankenau et al., 1994). We find an imperfect match to this consensus sequence 40 bp upstream of the transcription start (Fig. 2C), which coincides in its position with the localization of this motif in the β2-tubulin gene. In addition, we found within the proximal promoter sequences good matches to the elements, that were proposed to be involved in the testis-specific transcription of *gdl* (Schulz et al., 1990) and *hsp26* (Glaser and Lis, 1990) genes (Fig. 2C). Finally, a perfect match to the TTGTCAAAT box is localized at -860 bp upstream of the transcription start of the histone H3 antisense transcript. This box was identified in the promoters of the Mst87F, *Mst84Da-d* genes (Kuhn et al., 1991) and in the promoter of the testis-expressed myosin heavy-chain isoform (Miedema et al., 1995). Copies of this motif were found several hundred base pairs upstream of the transcription start site of these genes. Although the significance of all the mentioned promoter elements for the regulation of the histone H3 antisense transcription remains to be tested, it seems likely that at least some of them contribute to it.

3.3. Histone H3 antisense transcripts are testis-specific

The results of the Northern blotting indicated that the histone H3 antisense transcripts were testis-specific. This was confirmed by PCR with an antisense histone H3 primer HAS4 (see Fig. 2A) and oligo(dT) on pools of cDNAs derived from the RNA of embryos, ovaries and fly carcasses (Fig. 3). We could detect no transcripts similar to the ones described above, in tissues other than testis.

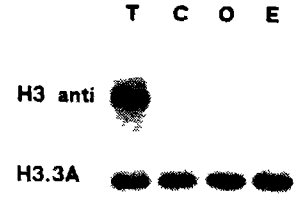


Fig. 3. Southern blot with products of PCR with the primers HAS4 and oligo(dT) on the pools of oligo(dT) primed cDNAs from *D. hydei* testes (T), carcasses (C), ovaries (O) and embryos (E) was hybridized to the histone H4 probe. As control PCR was performed on the same pools of cDNAs with primer H3-IV and oligo(dT) and hybridized to the *D. hydei* histone H3.3A 3' UTR probe (Akhmanova et al., 1995).

Using the same strategy, we could find no significant amount of histone H4 antisense transcripts, ending downstream of the H3 gene. One should bear in mind, that histone H3/H4 antisense transcripts, lacking the poly(A)-tail or heterogeneous in their 3' ends, could have avoided our detection by PCR. However, due to the fact that Northern blot analysis provides no indication for the presence of such transcripts, they are either absent or present in low abundance.

Our Northern blot results indicate, that there are no abundant testis-specific antisense transcripts of histones H1, H2A and H2B (not shown).

3.4. Histone H3 antisense transcripts are localized in the nuclei of primary spermatocytes

We have adapted the in situ hybridization protocol, developed for tissue sections (Dijkman et al., 1995), for the squashes of *Drosophila* testis. This protocol, in contrast to the previously published modifications of the proteinase K-based procedure used for *Drosophila* embryos (Tautz and Pfeifle, 1989), allows reliable detection of both nuclear and cytoplasmic transcripts in the male germ line cells and has the advantage of a better resolution than whole mount testis in situ hybridization. To avoid problems posed by crossreactions of in vitro transcribed RNA probes (see above), we used as probes mixtures of long sense or antisense oligonucleotides.

The antisense histone H3 probe hybridized strongly to a limited number of spermatogonial cysts (see Fig. 4.1A). This was in accordance with our expectations, because histone sense transcripts should be present in the mitotically dividing gonial cells. Due to the fact, that the histone expression is linked to the DNA replication, only those cells, which are in the S phase are expected to have large amounts of histone mRNA. After the completion of the DNA replication, the histone messengers are destabilised and disappear quickly (Schümperli, 1988; Marzluff and Pandey, 1988). All spermatogonia within one cyst are connected by cytoplasmic bridges and proceed through the cell cycle synchronously (Lindsley, 1980). That explains why only gonial cysts in the S phase display strong hybridization to the antisense histone probe, while neighbouring non-replicating cysts do not react.

The antisense histone H3 probe did not react either with nuclei or the cytoplasm of the primary spermatocytes (Fig. 4.1B). However, very early spermatocytes, which undergo DNA replication, are very small in size and are difficult to distinguish from spermatogonia after squashing and subsequent treatment used.

The sense probe displayed no reaction in the testis tip, but it hybridized to the nuclei of primary spermatocytes (Fig. 4.2A and 4.2B). The label was distributed throughout the nucleus and showed no clear association with any of the lampbrush loops, formed by the Y chromosome. The histone H3 antisense transcripts were most abundant in the stage II spermatocytes. Their level decreased rapidly during meiosis and they were undetectable in postmeiotic stages.

To confirm the specificity of our in situ hybridization procedure we used as

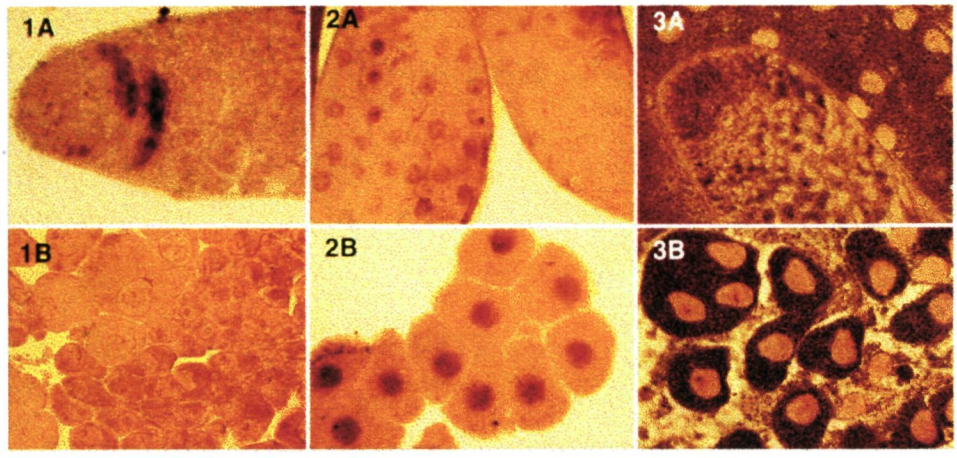


Fig. 4. In situ hybridization to testes squashes. The following probes were used: **1A** and **1B**, histone H3 antisense probe (mixture of the HAS1-HAS4 primers); **2A** and **2B**, histone H3 sense probe (mixture of the HS1-HS4 primers); **3A** and **3B**, histone H3.3A 3' UTR antisense probe (the same results were obtained with histone H3.3A 5'UTR-specific oligonucleotide). **1A** and **3A** - testis tip with spermatogonia and primary spermatocytes; **2A** - testis tube with primary spermatocytes **1B**, **2B** and **3B** - primary spermatocytes.

a control antisense oligonucleotides, derived from the 5' and 3' UTRs of the *D. hydei* histone H3.3A, which is strongly expressed in testes (Fig. 1B). Both oligonucleotides, which were used separately, hybridized strongly to the cytoplasm of all spermatogonia (Fig. 4.3A) and of the primary spermatocytes (Fig. 4.3B) The fact that all spermatogonia displayed signals is not surprising, since histone H3.3A is a replacement variant gene and its expression is not linked to the cell cycle.

3.5. Possible functions of the histone H3 antisense transcripts

The results of Northern blot hybridization, cDNA cloning by RT-PCR and in situ hybridization presented in this study are consistent with each other and, taken together, they prove the presence of histone H3 antisense transcripts in *D. hydei* male germ cells.

To assess further the question of a functional significance of histone H3 antisense transcription we have also analysed histone H3 gene expression in an evolutionary distant *Drosophila* species, *D. melanogaster*. By performing PCR on *D. melanogaster* cDNA we find male-specific transcripts, spanning the 5' part of histone H3 and the entire histone H4 gene (not shown). Such transcripts are, however, much less abundant in *D. melanogaster* compared to *D. hydei* and could not be detected on Northern blots of *D. melanogaster* RNA. Such a difference in the abundance of antisense transcripts might be related to the much

smaller size of male germ cells in *D. melanogaster*, their shorter developmental period or to species-specific differences in the accumulation and stability of these transcripts.

Sequence analysis of the *D. hydei* histone H3 antisense transcript has shown that the only long ORF, present in it, is that of the histone H4 protein. It is, however, preceded by multiple start and stop codons, which create highly unfavorable situation for the initiation of translation (Kozak, 1991). The longest ORF in the 5' part of the transcript consists of 40 codons and the encoded polypeptide does not display significant similarity to any proteins in sequence databases. Due to the fact that the antisense transcripts are not exported from the nucleus into the cytoplasm (see Fig. 4), it is highly unlikely that they are translated.

One has to conclude, therefore, that the function of the antisense transcripts is a regulatory one. Antisense RNAs were shown to regulate a variety of processes in prokaryotes (Wagner and Simons, 1994). Also in eukaryotes naturally occurring antisense transcription was observed for a number of genes (Kindy et al., 1987; Kimelman and Kirschner, 1989; Miyajima et al., 1989; Khochbin and Lawrence, 1989; van Duin et al., 1989; Hildebrandt and Nellen, 1992; Vellard et al., 1992; Dolnick, 1993; Laabi et al., 1994; Hsieh Li et al., 1995; Bedford et al., 1995; Murashov and Wolgemuth, 1996). However, only in a few cases their regulatory role was proven (Kimelman and Kirschner, 1989; Krystal et al., 1990; Munroe and Lazar, 1991; Hildebrandt and Nellen, 1992). It is possible that the antisense transcripts represent components of the mechanism for silencing of the cell-cycle regulated histone H3 genes in primary spermatocytes.

During the primary spermatocyte development DNA synthesis occurs very early (Hennig, 1967), and one would expect the cell-cycle regulated mRNAs to be quickly degraded afterwards, similar to what has been observed in other systems (Schümperli, 1988; Marzluff and Pandey, 1988). However, male germ line cells represent a special situation. Both in *D. melanogaster* and in *D. hydei*, normal polyadenylation signals are found downstream of the inverted repeat sequences of the cell-cycle regulated histone genes (Matsuo and Yamazaki, 1989; Kremer and Hennig, 1990). These polyadenylation signals are utilized to generate polyadenylated mRNAs, which, at least in case of the *D. melanogaster* H2B and H4 genes are most abundant in the testis (Akhmanova, unpublished). Such polyadenylated cell-cycle histone messengers are more stable than their non-polyadenylated counterparts (Kirsh et al., 1989; Nagata et al., 1991) and therefore are likely to persist during spermatocyte development. Several polyadenylation signals are present downstream of the *D. hydei* replication-dependent H3 gene. However, we could observe in the *D. hydei* testis RNA preparations no polyadenylated histone H3 mRNAs, resulting from the utilization of these signals. This could be due to the histone H3 gene repression by the antisense mechanism.

It is noteworthy that the histone H3.3A gene remains active in the spermatocytes (see above). An intriguing possibility is that the repression of the cell-cycle regulated H3 gene is necessary to achieve the substitution of the cell-

cycle regulated H3 histone by the H3.3 variant in male germ cells. Such a substitution might be important either for proper gene activation in spermatocytes or for the subsequent chromatin condensation-decondensation cycles (Kremer et al., 1986). The fact that we could find no significant histone H4 antisense transcription correlates with the absence of protein variants of histone H4 in *Drosophila*. We have shown that the cell-cycle regulated and the recently discovered replacement histone H4 genes encode exactly the same protein (Akhmanova et al., 1996). Hence, a substitution of the H4 protein is not required.

One can imagine several mechanisms for the antisense-mediated histone H3 gene repression. Antisense transcription could impair sense transcription by interfering with the RNA polymerase progression along the opposite strand (Elledge and Davis, 1989). It could also inhibit transcription initiation from H3 and H4 promoters by preventing the stable binding of transcription factors to the H3-H4 intergenic region. Finally, the antisense RNA could form duplexes with the sense RNA and interfere with its processing, stability or nuclear export. Duplex formation between sense and antisense transcripts from the same gene was implied in the regulation of the pre-mRNA splicing (Krystal et al., 1990, Munroe and Lazar, 1991), stability (Hildebrandt and Nellen, 1992) or conversion of adenine residues to inosine (Kimelman and Kirshner, 1989). The fact that we could not detect histone H3 sense transcripts in spermatocytes does not necessarily contradict the possibility of duplex formation. In several cases where the regulatory function of the antisense RNAs was proven, the latter were found in a considerable excess over the sense RNA molecules (Kimelman and Kirshner, 1989; Hildebrandt and Nellen, 1992)). Sense histone H3 mRNA may therefore be low in abundance and escape our detection.

The antisense histone H3 RNA represents one of the few examples of naturally occuring antisense transcription in *Drosophila* (Lankenau et al., 1994; Henikoff et al., 1986; Spencer et al., 1986; Chen et al., 1987; Contursi et al., 1993). In most of such cases antisense transcription results from a complex gene arrangement, for example where genes either overlap in their 3' regions (Spencer et al., 1986) or are localized in an opposite orientation in introns of another gene (Henikoff et al, 1986; Chen et al., 1987). Non-coding antisense transcripts with putative regulatory functions were found in two types of retrotransposons: the F elements (Contursi et al., 1993) and the *micropia* elements (Lankenau et al., 1994). Remarkably, also the *micropia* antisense RNA is testis-specific. This suggests that antisense transcription is not uncommon in *Drosophila* male germ line cells. It should be noted, however, that the *micropia* antisense RNAs are preferentially localised in the cytoplasm, while the histone H3 antisense RNA is found in the nuclei of the primary spermatocytes (Lankenau et al., 1994). Therefore, if both antisense transcripts are involved in gene regulation, the molecular mechanisms of their action may be different.

Acknowledgements. We thank Petra Bindels, Harry Harhangi and Wenli Gu for help with conducting the experiments and valuable discussions. This research was supported by the European Community Human Capital Programme (Networks), contract No. CHRX-CT94-0511.

Chapter VI

Characterization of an antiserum, specific for the histone H3.3 replacement variant, and analysis of the histone H3.3 distribution in *Drosophila* male germ line cells

Anna Akhmanova[1], Koos Miedema[1], Yaxian Wang[2], Mieke van Bruggen[3], Sylviane Muller[4], Jo H.M. Berden[3], Evangelos N. Moudrianakis[5] and Wolfgang Hennig[1,2]

[1] Department of Molecular and Developmental Genetics, Faculty of Sciences, Catholic University of Nijmegen, Nijmegen, The Netherlands

[2] Max Plank Guest Laboratory at the Institute of Cell Biology, Academia Sinica, Shanghai, China

[3] Division of Nephrology, University Hospital, Nijmegen, The Netherlands

[4] Institut de Biologie Moleculaire et Cellulaire, CNRS, Strasbourg, France

[5] Department of Biology, The Johns Hopkins University, Baltimore, USA

Abstract

A rabbit antiserum, specific for the histone H3.3 replacement variant, has been raised. An H3.3 peptide, derived from the histone H3 sequence region 83-91, where most of the differences between the repacement histone H3.3 protein and the cell-cycle regulated variants H3.1 and H3.2 are localized, was used for the immunization. Western blot experiments have shown the specificity of this polyclonal antibody for the H3.3 histone variant protein. In addition, we have shown on Western blots that two antinuclear monoclonal antibodies (MAbs) display strong reactivity with the H3.3 histone, but not with its replication-dependent counterparts. These MAbs were isolated from mice with systemic lupus erythematosus (SLE), an autoimmune disease which is characterized by the presence of anti-DNA and anti-nucleosome antibodies. This indicates that the H3.3 histone might play a role as autoantigen in SLE.

We have applied the antiserum, raised against the H3.3 peptide, to cytological preparations of *Drosophila* testes, because our previous studies have shown that a histone H3.3-encoding gene is strongly expressed in *Drosophila* male germ line cells. The antiserum reacted with some of the lampbrush loops, formed by the Y chromosome during the meiotic prophase and with chromatin of the post-meiotic male germ cells. Our data indicate that histone H3.3 is not randomly distributed, but concentrated in certain parts of chromatin. The H3.3 histone is removed from the spermatid nuclei along with the other core histones during the late stages of spermatogenesis. In *Drosophila* polytene chromosomes a rather uniform distribution of the histone H3.3 was observed.

1. Introduction

Histones are the elementary constituents of the nucleosome, which is the structural unit of chromatin. In higher eukaryotes, histone genes can be divided into three categories, replication-dependent, replacement variant, and tissue-specific variant genes. Most of the histone replacement variant genes differ from their replication-dependent counterparts not only by their mode of regulation, copy number, gene structure and the type of mRNA they synthesize, but also in the amino acid sequence of the encoded proteins. For example, the animal histone H3 replacement variant, histone H3.3, differs by several amino acid substitutions from the replication-dependent histones H3 (see Fig. 1) and these differences are highly conserved throughout the animal kingdom (Brush et al., 1985; Wells and Kedes, 1985; Wellman et al., 1987; Swenson et al., 1987; Fretzin et al., 1991; Fucci et al., 1994). The evolutionary conservation of the histones H3.3 suggests that their protein structure is of functional significance. However, since the replacement H3 variants have apparently arisen independently in plants, animals and *Tetrahymena* (Thatcher et al., 1994), it was argued that only the replication independence of gene expression is functionally important. The differences in the protein structure were suggested to be selectively neutral polymorphisms.

At present, little direct data on the function of the replacement variants of

the histone H3 and their chromosomal distribution are available. In a number of studies the H3.3 protein synthesis was analysed in relationship to the cell-cycle phase. A classical example is the work of Wu et al. (1982), where it was shown that H3.3 histone is the only histone H3 variant synthesized during G1, G2 and G_0 stages of Chinese hamster ovary cells. Newly synthesized histone H3.3 is preferentially found in the chromatin fraction, enriched in transcriptionally active or competent chromatin of non-replicating cells (Hendzel and Davie, 1990). In differentiated cells histone H3.3 accumulates with time and gradually becomes the predominant H3 subtype (Zweidler, 1984).

The most common way to analyse the distribution of a particular protein within a cell is by applying antibodies, specific to this protein. This approach was successfully used for studying the variants of histone H2A (Stargell et al. 1993) and histone H1 (Mohr et al., 1989) and for analysing the distribution of differentially acetylated isoforms of histone H4 (Turner et al., 1992; for review see Turner, 1993). We have addressed the question whether it is possible to distinguish immunologically between the histone H3.3 variant and its replication-dependent counterparts. Due to the very small number of differences between the H3 variants and the fact that most of these differences are clustered within a rather hydrophobic part of the H3 protein (amino acid positions 87-90), such a possibility is not directly evident. We have found that an antiserum, raised against the histone H3.3 peptide, corresponding to amino acid positions 83-91 (anti-H3.3p antiserum), can indeed recognize specifically the H3.3 variant protein. In addition, we have also found that certain mouse auto-antibodies can also distinguish between replication-dependent and independent histone H3 variants.

To test whether the anti-H3.3p antiserum can be applied for studying chromatin, we have used *Drosophila* male germ line cells and polytene chromosomes. Histone H3.3 distribution in the male germ cell chromatin was of a particular interest to us, since we have shown previously, that one of the two *Drosophila* histone H3.3-encoding genes, H3.3A, is strongly expressed in testes (Akhmanova et al., 1995). Also, at least for *D. hydei*, quite detailed information is available on the organisation and rearrangements of chromatin in the developing male germ line cells (Kremer et al., 1986). We demonstrate that the anti-H3.3 antiserum can indeed react with the chromatin, provided that the DNA is at least partially removed. This observation is agreement with the data on the nucleosome structure, which shows that the part of the histone H3, recognized by our antiserum (residues 83-91), is lying under the path of the DNA (Arents et al., 1991; Arents and Moudrianakis, 1993; this paper). Non-uniform distribution of the histone H3.3 in the chromatin of the male germ cells was observed. The implications of our findings are discussed with respect to chromatin structure and function in male germ cells.

The polytene chromosomes from salivary glands reacted with the anti-H3.3p antiserum in a rather uniform way, providing no clues about the possible functions of the H3.3 histone variant.

2. Materials and Methods

2.1. Preparation of antisera, affinity purification and antibody reactions

The H3.1/2 peptide with the sequence RFQSSAVMAC and the H3.3 peptide with the sequence RFQSAAIGAC (see Fig. 1) were synthesized in the laboratory of Dr. J. Zimmermann, Jena. The C-terminal cysteine residue is not present in the histone H3 sequence and was introduced for coupling to the carrier. The peptides were coupled to maleimide-activated keyhole limpet hemocyanin (KLH) (Boehringer Mannheim) according to the instructions of the supplier. Immunization of rabbits was performed with 500 µg of the KLH-coupled peptide in Freunds complete adjuvant (Harlow and Lane, 1988). Rabbits received three injections with intervals of four weeks.

The polyclonal antibodies were affinity purified over the peptide H3.3 coupled to a Sulfolink column (Pierce). After washing with PBS (phosphate-buffered saline, containing 137 mM NaCl, 27 mM KCl, 1.5 mM KH_2PO_4 and 6.5 mM Na_2HPO_4 (pH 7.4)) and 0.5 M NaCl/10 mM Tris-HCl pH7.5, H3.3 specific antibodies were eluted with 0.1 M glycine, pH 2.6 (Harlow and Lane, 1988). Eluted antibodies were immediately neutralized with 1 M Tris-HCl pH 9.5. Bovine serum albumine (BSA) was added to a final concentration of 1% and the antibodies were either used directly on blots or concentrated with a Centricon 30 device (Amicon) prior to immunocytology.

The polyclonal anti-H3 serum (Termaat et al., 1992) was diluted 1 : 500 (PBS/1% BSA) for Western blotting and 1:50 for immunocytology. The polyclonal antibodies against histone H2A (serum 1 in Muller et al., 1986), histone H2B (serum 1 induced in the presence of RNA in Muller et al., 1991), histone H3 (serum 1 induced in the presence of RNA in Muller et al., 1991) and anti-H4 (serum 2 induced in the presence of RNA in Muller et al., 1991) were diluted 1 : 100 (PBS/1% BSA) and 1:10 for immunocytology. The monoclonal antibodies #42, #56 (Smeenk et al., 1988; Stemmer et al., 1996) and KM2 (Kramers, 1995; van Bruggen et al., 1996) were used as undiluted culture supernatants. Incubations were performed for 1 hr at room temperature or overnight at $4^{O}C$. Secondary goat anti-rabbit or goat anti-mouse antibodies were diluted at 1 : 2500 and the blots were incubated for 30-60 min. The affinity purified anti-H3.3p antibodies were tested on PVDF blots spotted with 1 µg of H3.3 and H3.1/2 peptides and H3 protein from calf thymus (Boehringer Mannheim). Dot blots were treated and incubated as described below for Western blots.

2.2. Separation of histones on acetic acid-urea-Triton (AUT) gels and Western blotting

The calf thymus histones (Boehringer Mannheim) were separated on 5% acetic acid/6M urea/0.4% Triton X-100 12% polyacrylamide gels (0.4 mm thick), essentially according to Lennox and Cohen (1989). Special care was taken to prevent oxidation of methionine-containing proteins during electrophoresis by a cysteamine scavenge step according to Lennox and Cohen (1989). After electrophoresis gels were either stained directly in 0.25% Coomassie Brilliant Blue R-250 (45%

methanol/10% acetic acid) for 20 min and destained in 45% methanol/10% acetic acid or placed directly for 2 hrs in 0.025 M Tris/0.192 M Glycine/1% SDS to equilibrate before Western blotting (Whitfield et al., 1986). Subsequently, gels were transferred to a semi-dry blot apparatus (Phase). Proteins were blotted to PVDF membrane (Millipore) for 45 min at 300 mA. After blotting, the membrane was blocked in 3% BSA/2% milkpowder/PBS for 1 hr at 37°C, air dried and stored at 4°C or washed in PBS prior to antisera incubations.

2.3. Indirect immuno-fluorescence cytology

2.3.1. Testis squashes

Testis squashes were prepared as described earlier (Kremer et al., 1986) and fixed in ethanol for 30 sec. After fixing for 10 min in 4% paraformaldehyde/PBS, slides were washed once for 5 min in PBS. Slides were incubated for 15 min in 1% Triton X-100 in PBS and washed twice in PBS. If a DNase treatment step was included, slides were washed twice in nick-translation buffer (50 mM Tris-HCl (pH 7.5), 10 mM $MgSO_4$, 50 µg/ml BSA) and incubated with 4% DNase I in nick-translation buffer for one hour at room temperature. After that the squashes were washed three times for 10 min in PBT (PBS containing 2 mg/ml of BSA and 0.1% Triton X-100). The incubation with the first antibody was performed overnight at 4°C and after washing 3 times for 10 min in PBT, the squashes were incubated for 1 hr with the secondary antibody (FITC-conjugated goat anti-rabbit Fab fragments, diluted 1:50 in PBT). The slides were washed again 3 times in PBT for 10 min and once in PBS and mounted in the *p*-phenylenediamine/glycerol mountant. The slides were examined with a Zeiss photomicroscope III with epifluorescence equipment. Photographs were taken on Kodak Ektachrome 400 film.

For chemical degradation of the DNA trichloracetic acid (TCA) treatment was used. Testis squashes were prepared and fixed with ethanol as described above. The slides were subsequently incubated in 4% paraformaldehyde/PBS at 4°C for 16-20 hrs, washed several times in water and incubated in 5% TCA at 60°C for 3 hrs (Hauschteck-Jungen and Hartl, 1982). After washing in 70% ethanol and in PBS slides were incubated for 1 hr in PBT. The squashes were washed in PBS and PBT and incubated with the first and second antibodies as described above.

2.3.2. Polytene chromosomes

Polytene chromosomes were prepared essentially as described by Ashburner (Protocol 30 in Ashburner, 1989) and treated as described for testis squashes (see above).

3. Results

3.1. Isolation of an antiserum specific for the histone H3.3 protein

The histone H3 variant proteins in different organisms are strongly conserved and vary only by a few amino acids. For example, the vertebrate H3.1 and H3.2 proteins differ by one amino acid (position 96, C→S) from each other (Fig. 1). The *Drosophila* cell-cycle regulated histone H3 corresponds in sequence to the

```
Bt H3.1    ARTKQTARKSTGGKAPRKQLATKAARKSAPATGGVKKPHRYRPGTVALREIRRYQKSTELLIRKLPFQ  68
Bt H3.2    ------------------------------------------------------------------
Dm H3cc    ------------------------------------------------------------------
   H3.3    ------------------------------S-----------------------------------

H3.1/2pep                RFQSSAVMAC
Bt H3.1    RLVREIAQDFKTDLRFQSSAVMALQEACEAYLVGLFEDTNLCAIHAKRVTIMPKDIQLARRIRGERA  135
Bt H3.2    ----------------------------S-------------------------------------
Dm H3cc    ----------------------------S-------------------------------------
   H3.3    -----------------A-IG-----S---------------------------------------
H3.3 pep                 RFQSAAIGAC
```

Fig. 1. Amino acid sequence alignment of the animal histone H3 variants. *Bos taurus* (Bt) H3.1 and H3.2, animal H3.3 and *D. melanogaster* (Dm) cell-cycle regulated H3 are shown. The original *D. melanogaster* histone H3 amino acid sequence (Matsuo and Yamazaki, 1989) contained an isoleucine (I) at position 117. We resequenced this region in several cDNA clones and found that it encoded valine (V) at position 117 in all clones, similar to the H3.1/2 and H3.3 sequences. The identical residues are indicated with dashes. The α-helical regions of the C-terminal domain are indicated with lines above and below the sequence (Arents et al., 1991; Wolffe and Pruss, 1996b). The sequences of the synthetic peptides, used in this study are shown above and below the alignment. A C-terminal cysteine, present in both peptides, is absent in the histone sequence. It was included to facilitate the coupling to the carrier.

vertebrate H3.2. The histone H3.3 variant is identical in all vertebrates and invertebrates studied so far. It differs from the histone H3.2 subtype only by four amino acid substitutions (positions 31, 87, 89, and 90, Fig. 1). In order to distinguish immunologically between the histone H3.1/2 and H3.3 proteins, we used the region 83 to 91 from the histone H3 amino acid sequence to synthesize two peptides, specific for the H3.1/2 and H3.3 variants. These peptides were used to raise polyclonal antibodies.

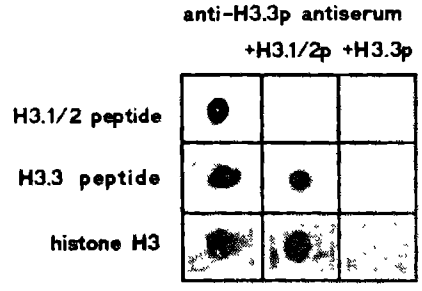


Fig. 2. Dot blots with 1 μg of H3.1/2 and H3.3 peptides and calf thymus histone H3 incubated with the affinity purified anti-H3.3p antibodies. Middle and right columns show the reaction of the anti-H3.3p antibodies, preincubated, respectively, with H3.1/2 and H3.3 peptides.

Two different antisera induced against the H3.1/2 peptide displayed very low reactivity with the histone H3 protein and were not analysed further. In contrast, the antibodies against the H3.3 peptide (anti-H3.3p antibodies) were quite reactive, especially after affinity purification over the H3.3 peptide. The immunoreactivity of the affinity purified anti-H3.3p antibodies with H3.1/2 and H3.3 peptides and the histone H3 fraction (from calf thymus) on dot blots is shown in Fig. 2. The anti-H3.3p antibodies react with the H3 protein and with

both peptides. The latter is not surprizing since only three out of nine amino acids in the two peptides are different. Nevertheless the reaction of the anti-H3.3p antibodies with the H3.1/2 peptide on the dot blot can be completely inhibited by the preincubation of the same antibodies with this peptide added as competitor in solution. Such preincubation hardly reduces the reaction of the antibodies with the H3.3 peptide and the histone H3 protein. However, if the H3.3 peptide is used for the preincubation, all the signals disappear. This shows that the anti-H3.3p antibodies specifically recognize the H3.3 peptide, when preincubated with the H3.1/2 peptide as competitor.

Characterization of the anti-H3.3p antibodies was extended by Western blot analysis of calf thymus histones separated on AUT gels. We used AUT gels with 6M urea and 0.4% Triton, which are optimal for separation of histone H3 variants (Zweidler, 1978). Histone mix from calf thymus is well suited for this analysis, because relative positions of all calf thymus histone variants in these separation conditions are well defined (Zweidler, 1984; Boulikas, 1985) and can be used as a reference. The affinity purified anti-H3.3p antibodies showed a specific reaction with a single major band which corresponded in mobility to the histone H3.3 variant (Fig. 3a). This reaction could be completely inhibited by the preincubation of the antibodies with the H3.3 peptide, but not with the H3.1/2 peptide (Fig 3a).

For comparison, we have used an antiserum raised against the whole histone H3 fraction from calf thymus (Termaat et al., 1992). This anti–H3 serum shows strong reaction with histone H3.1 variant, which is the most abundant H3 subtype in calf thymus histone mix (Fig.3a, lane aH3). The band in the upper part of the gel (marked with an asterisk in Fig. 3) probably corresponds to a histone H3 dimer, judging from its mobility and the fact that all sera with affinity to the histone H3 react with it (see below). The anti-H3 serum is not completely specific for histone H3, since a number of additional weak bands are present on Western blots. It apparently crossreacts with histone H2B and histone H2A.Z (Fig. 3a).

The serum against total H3 fraction shows only a weak signal at the level of the H3.3 variant (Fig. 3a and b). This reflects the low concentration of this histone in the calf thymus histone mix. However, this signal was considerably increased, when the serum was affinity purified over the H3.3 peptide (Fig. 3b, lane aH3ap). At the same time, the reaction with the H3.1 variant was reduced, although not completely abolished, while the signals corresponding to other histone types, such as H2B, disappeared completely. This supports the correctness of our assignment of the position of the H3.3 variant on the Western blot.

The positions of other core histones on the blots were confirmed with the help of antisera raised against individual chicken histones H2A, H2B, H3 (marked with aH3*), and H4 (Muller et al. 1986; Muller et al., 1991; Briand et al. 1992; see Fig. 3b). Due to the extremely high degree of conservation of core histones these sera where expected to crossreact with calf thymus histones. All four sera reacted with proteins of the expected mobility on the AUT gel, which was

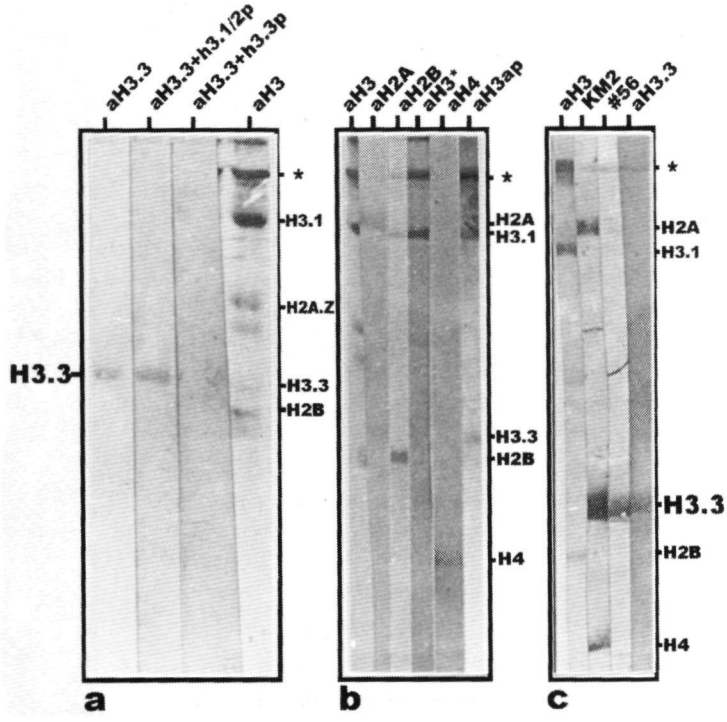


Fig. 3. Western blots from calf thymus histones, separated on AUT gels, incubated with different anti-histone antisera. Panels **a**, **b** and **c** represent blot stripes from three different AUT gels. The following incubations were performed: **aH3.3** - anti-H3.3p antibodies ; **aH3.3+h3.1/2p** - anti-H3.3p antibodies, preincubated with H3.1/2 peptide; **aH3.3+h3.3p** - anti-H3.3p antibodies, preincubated with H3.3 peptide, **aH3** - anti-H3 serum, raised against calf thymus histone H3 fraction; **aH2A** - anti-H2A serum, raised against chicken histone H2A; **aH2B** - anti-H2B serum, raised against chicken histone H2B; **aH3*** - anti-H3 serum, raised against chicken histone H3; **aH4** - anti-H4 serum, raised against chicken histone H4; **aH3ap** - anti-H3 antibody, raised against calf thymus histone H3, affinity purified over the H3.3 peptide; **KM2** and **#56** - monoclonal antibodies from SLE mice. The band, indicated with an asterisk, is probably an H3 dimer.

clearly distinct from the mobility of the H3.3 variant. It should be noted that the anti-H2B serum recognized not only the histone H2B, but also histone H3.1 (and probably its dimer).

Taken together, our data demonstrate that the affinity purified polyclonal anti-H3.3p antibodies recognize specifically the histone H3.3 variant.

3.2. Mouse auto-antibodies with high affinity to the histone H3.3

The presence of auto-antibodies, reacting with DNA, histones or whole nucleosomes is one of the characteristic manifestations in a number of autoimmune

diseases, among them SLE (Monestier and Kotzin, 1992; Kramers et al., 1993). Our interest in these auto-antibodies was stimulated by the finding that some MAbs, isolated from lupus mouse strains, when tested in ELISA, displayed strong reaction with the peptide 83-100 of the histone H3, but not with any other peptides, covering the entire amino acid sequence of histone H3 and other core histones (Kramers, 1995; Stemmer et al., 1996). Since the region, where most of the differences between variants of animal histone H3 are localized, is situated between amino acids 83 and 100 (see Fig. 1), we were interested if these antibodies were specific for one of the H3 variants. From the two MAbs analysed, one (#42) did not react with any histone proteins on Western blot (not shown), while the other one (#56) reacted specifically with the H3.3 variant (Fig. 3c). In addition, we found another antinuclear MAb, KM2, which displayed strong reactivity with the histone H3.3 but not with the H3.1 variant (Fig. 3c). This antibody also recognized histones H2A and H4, which is in agreement with the data on characterization of this MAb by ELISA (van Bruggen et al., 1996).

3.3. Immunocytology with the anti-H3.3p antibodies

The structure of the octameric histone core of the nucleosome has been determined to a resolution of 3.1 Å (Arents et al., 1991). According to this structure, the region 87-90, where the histone H3.3 differs from the histones H3.1/2, is localized at the N-terminal end of the long α-helix (helix 3) of the histone H3 C-terminal domain (Arents et al. 1991; see also Fig. 1). The residues 86, 87, 89 and 90 are all clustered together on the surface of the histone octamer (see Fig. 4a). According to the model of the nucleosome (Arents and Moudrianakis, 1993), these residues are covered by the DNA double helix, which is wrapped around the octamer (Fig. 4b). The presence of the DNA, therefore, is likely to make the histone H3.3-specific epitope inaccessible for antibodies.

Since the H3.3 variant is conserved in all animals, the anti-H3.3p antibodies can be applied to any animal system. We have chosen to test the anti-H3.3p antibodies on cytological preparations of *Drosophila* tissues, because in our earlier studies we have characterized the expression of the histone H3.3 genes in this organism on the RNA level (Akhmanova et al., 1995). In addition, analysis of the histone H3.3 distribution in *Drosophila* is necessary to form a basis for subsequent genetic studies of the function of this protein.

When we used standard fixation protocols, which leave most of the DNA intact, the anti-H3.3p antibodies were completely unreactive (not shown). Therefore, special effort has been taken to remove the DNA, either enzymatically (with high concentrations of DNase I) or chemically (with hot TCA). As expected, the degradation of the DNA led to a considerable increase in the stength of binding of the anti-H3.3p antibodies. In all cases control preincubations of the anti-H3.3p antibodies with H3.1/2 and H3.3 peptides were performed. The preincubation with the H3.1/2 peptide had little impact on the binding of the antibodies, while the preincubation with the H3.3 peptide almost totally abolished it (an example of such a preincubation experiment for the spermatocyte nuclei is shown below).

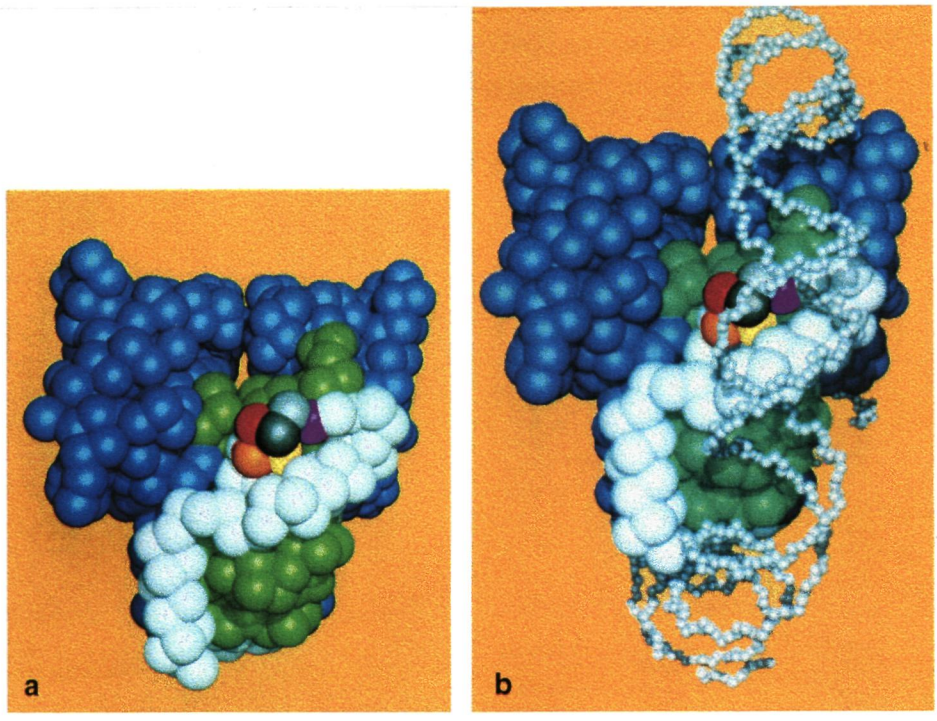


Fig. 4. Positions of the histone H3.3-specific substitutions in the histone octamer. **a.** The view of the histone octamer as it appears by looking down at a plane which contains the dyad and the superhelical axes. The dyad axis runs from top to bottom. The H2A-H2B dimers are shown in blue, histone H3 in white and histone H4 in green. Residue 84 of the histone H3 is shown in purple, 85 is light grey, 86 is dark grey, 87 is red, 89 is yellow and 90 is orange.
b. The same view of the octamer, as in **a**. The backbone of the DNA double helix, wrapped over one half of the histone octamer, is shown.

3.4. Polytene chromosomes

Treatment of the preparations of *D. melanogaster* polytene chromosomes with DNase I and with TCA gave very similar results. In both cases, a rather general staining of the chromosomes was observed. Neither the heterochromatic regions (the chromocenter) nor the highly transcribed regions (the puffs) showed preferential staining with the anti-H3.3p antibodies (Fig. 5). The intensity of the antibody reaction roughly followed the DNA concentration, being somewhat higher in the bands of the polytene chromosomes. The signals were not very strong, indicating that the H3.3 variant represents a minor component of the polytene chromosomes.

3.5. Male germ cells

We were particularly interested in the histone H3.3 distribution in the chromatin

of the *Drosophila* male germ cells, because we have shown earlier, that the histone H3 3A gene is strongly expressed in *Drosophila* testes (Akhmanova et al., 1995) The process of spermatogenesis was described in detail both in *D melanogaster* (Lindsley and Tokuyasu 1980) and in *D hydei* (Hennig and Kremer, 1990) In addition the distribution of the DNA at different stages of male germ cell development was analysed in *D hydei* (Kremer et al., 1986) and will be used here as a reference

3 5 1 Premeiotic stages

Spermatogenesis starts with the division of the stem cells giving rise to primary spermatogonia, which then undergo several mitotic divisions Spermatogonia are small round cells which fill the apex of *Drosophila* testis Using the anti-H3 3p antibodies on TCA or DNase I-treated testes squashes, we could detect no signal in the nuclei of spermatogonia (Fig 6a) In contrast, the antisera against total histone H3 (Fig 6b), as well as antisera against other core histones (not shown) strongly labeled the spermatogonial nuclei The distribution of the label was similar to that observed earlier for the DNA (Fig 1 in Kremer et al 1986)

After the male germ cells enter the meiotic profase (primary spermatocyte stage) they start to enlarge and become highly transcriptionally active One of the manifestations of this transcriptional activity is the unfolding of the Y-chromosomal lampbrush loops which fill most of the nuclear volume (for review see Hennig 1987) The condensed autosomes and the X-chromosome are localised at the nuclear periphery (Fig 2 of Kremer et al., 1986)

Staining of the TCA-treated primary spermatocytes of *D melanogaster* with the anti-H3 3p antibodies is shown in Fig 6c One or two strongly fluorescent foci per nucleus were observed This immunofluorescence pattern is clearly distinct from that obtained with the similarly treated spermatocytes after staining with the serum against total histone H3 (Fig 6e) In TCA-treated cells the anti-H3 serum (as well as the sera against H2A H2B and H4 not shown) labels mainly the condensed chromatin The distribution of the label corresponds to the DNA distribution as judged from DAPI staining (Fig 2 in Kremer et al., 1986) Comparison of the anti-H3 and anti-H3 3p patterns indicates that in the most of the condensed chromatin the concentration of the H3 3 histone variant is low It contains apparently predominantly the cell-cycle regulated H3 subtype

The decondensed chromatin gave no reaction in the TCA-treated preparations We thought that this might result from the complete loss of decondensed chromatin fibrils due to extensive hydrolysis in hot TCA Therefore, we applied to testis squashes a milder treatment with DNase I After such treatment strong signals in the nuclei of the *D melanogaster* spermatocytes were obtained (Fig 7a, b) The distribution of the label partially coincided with the localization of some of the Y-chromosomal lampbrush loops. Similar signals were obtained with the total H3 serum (not shown)

Due to the large size, characteristic position and morphology, the assignment of particular Y-chromosomal lampbrush loops is considerably easier in *D. hydei*, than in *D. melanogaster*. After staining of the DNase I-treated primary spermatocytes of *D. hydei* with the anti-H3.3p antibodies, the strongest reaction was obtained with the loop pairs pseudonucleolus (Ps) and clubs (Cl) (Fig. 7**d**, **e**). The reactions could be completely inhibited by the preincubation with the H3.3 peptide (Fig. 7**g**), but not by the preincubation with the H3.1/2 peptide (Fig. 7**f**). Similar labeling patterns were obtained when the *D. hydei* spermatocytes were stained with sera against other core histones, H2B and H4 (Fig. 8**a**,**b**). It should be noted that in the absence of TCA treatment the condensed chromatin was almost unreactive with all anti-histone sera. The nucleolus was also devoid of label, irrespective of the fixation protocol.

3.5.2. Postmeiotic stages

Meiotic chromosomes, due to their highly condensed nature, are not suited for immunofluorescence studies of the distribution of the histone variants (not shown).

Staining in postmeiotic male germ cells with anti-H3.3p antibodies could only be obtained after a TCA treatment of the tissue. This correlates with the absense of transcription and high chromatin compaction in *Drosophila* spermatogenic cells after meiosis. The patterns, obtained in *D. melanogaster* and *D. hydei* were very similar. The early spermatid nuclei could be best observed in *D. hydei*, due to their larger size and more clear morphology (Fig. 9). In early spermatids a large round structure, called a protein body, is present in the middle of the nucleus (Kremer et al., 1986). The anti-H3.3p antibodies gave strong labelling of this structure (marked with Pb in Fig. 9). In addition, labeling of the chromatin was also observed (Fig. 9**a**). When the spermatids were stained with the antibodies against total H3 protein, the main signal was observed in the chromatin, with protein body showing only very weak fluorescence (Fig. 9**c**). This indicates that most of the histone H3 in chromatin is represented by the cell-cycle regulated H3 variant. The concentration of histone H3 in the protein body is apparently low when compared to the chromatin and it is represented by the H3.3 histone.

During spermatid elongation, some of the histone H3.3 protein remains bound to the chromatin. In addition to the protein body, the chromatin lying under the membrane on the one side of the elongating nucleus is strongly stained (marked with M in Fig. 9**e**). The H3.3 histone is apparently not evenly distributed in the chromatin. This is clear from the comparison of the staining with anti-H3.3p and anti-H3 sera. The latter gives a pattern very similar to that obtained for the DNA with DAPI (Fig. 7 in Kremer et al., 1986), with the label distributed throughout the elongating nucleus, and not concentrated on one side of it (Fig. 9**g**, **h**).

In later elongation stages no major differences were observed between the staining with anti-H3.3p and anti-H3 antibodies. It should be noted that the H3.3-specific signal appeared to be more dispersed, than that obtained with the

Fig 5 Distribution of the histone H3 3 in the TCA-treated polytene chromosomes **a** Immunofluorescence of the polytene chromosomes from third instar larvae after incubation with the anti-H3 3p antibodies **b** Enlargement of a part of the polytene chromosome, demonstrating the uniform staining pattern

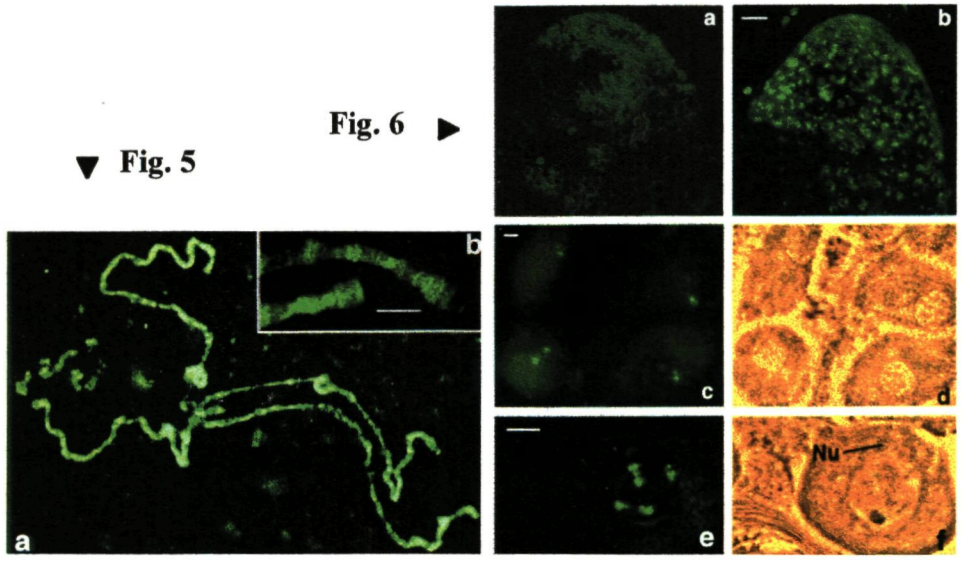
Fig 6 Pre-meiotic male germ cells of *D melanogaster* after TCA treatment of the tissue **a** Immunofluorescence of the testis tip with spermatogonia after the incubation with anti-H3 3p antibodies The background was enhanced to show the absense of nuclear signals **b** Immunofluorescence of the testis tip with spermatogonia after the incubation with anti-H3 serum The nuclei of spermatogonia are strongly labeled Bar represents 50 μm
c,d Primary spermatocytes after incubation with anti-H3 3p antibodies **c** Immunofluorescence **d** Phase contrast The reaction is observed in one or two foci in the nucleus Bar represents 5 μm
e, f Primary spermatocytes after incubation with anti-H3 serum **e** Immunofluorescence **f** Phase contrast Condensed autosomes and the X-chromosome react with the serum **Nu** - nucleolus Bar represents 5 μm

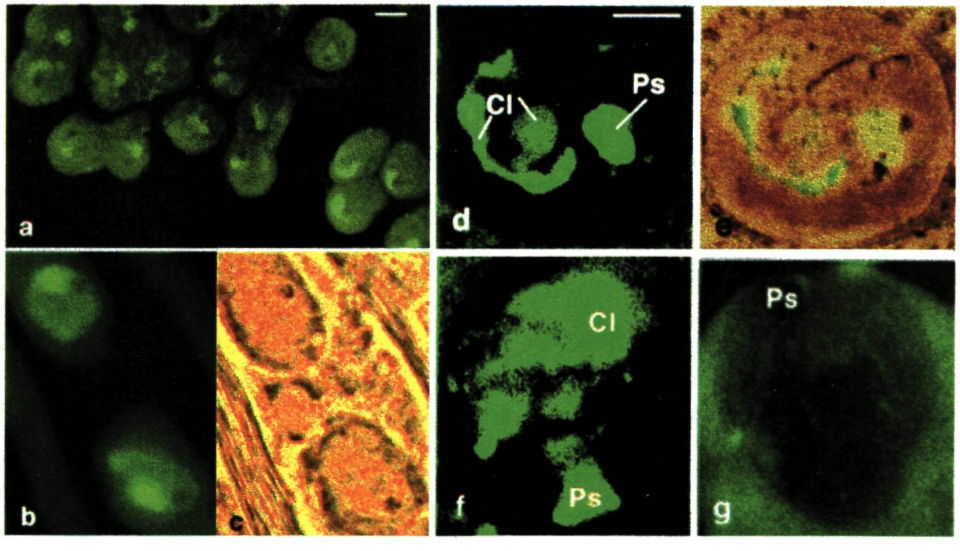
Fig 7 Primary spermatocytes after DNase I treatment **a, b, c** *D melanogaster* primary spermatocytes after the incubation with the anti-H3 3p antibodies **a, b** Immunoflourescence **c** Phase contrast
d, e, f, g Enlarged nuclei of the *D hydei* primary spermatocytes **d** Immunofluorescence after the incubation with the anti-H3 3p antibodies **e** Phase contrast image of the same nucleus **f** Immunofluorescence after the incubation with the anti-H3 3p antibodies preincubated with the H3 1/2 peptide **g** Immunofluorescence after the incubation with the anti-H3 3p antibodies, preincubated with the H3 3 peptide **Ps** - pseudonucleolus, **Cl** - clubs Bar represents 10 μm

Fig 8 Nuclei of the *D hydei* primary spermatocytes after DNase I treatment **a** Immunofluorescence after the incubation with the anti-H2B serum **b** Immunofluorescence after the incubation with the anti-H4 serum **Ps** - pseudonucleolus, **Cl** - clubs Bar represents 10 μm

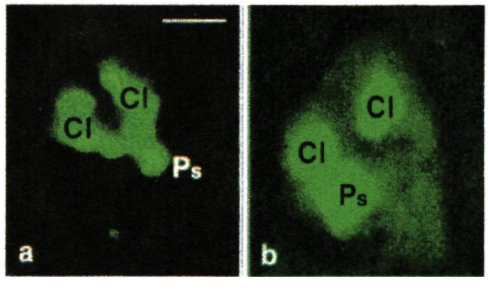
▼ Fig. 5

Fig. 6 ▶

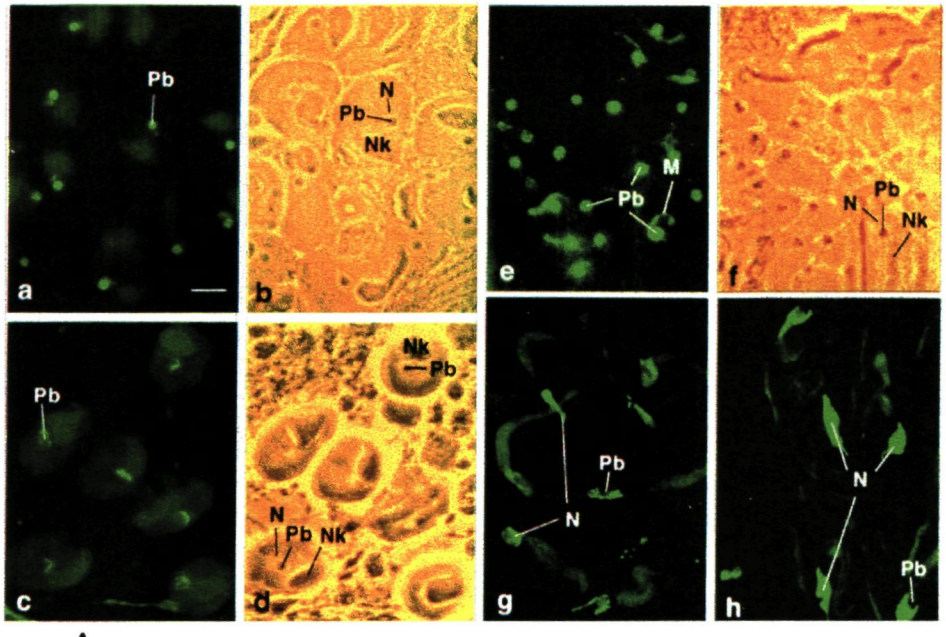




▲ Fig. 7



◀ Fig. 8


Pb
a
N
Pb
Nk
b
M
Pb
e
Pb
N
Nk
f
Pb
c
Nk
Pb
N
Pb
Nk
d
Pb
N
g
N
Pb
h

▲ Fig. 9

Fig. 10 ▼

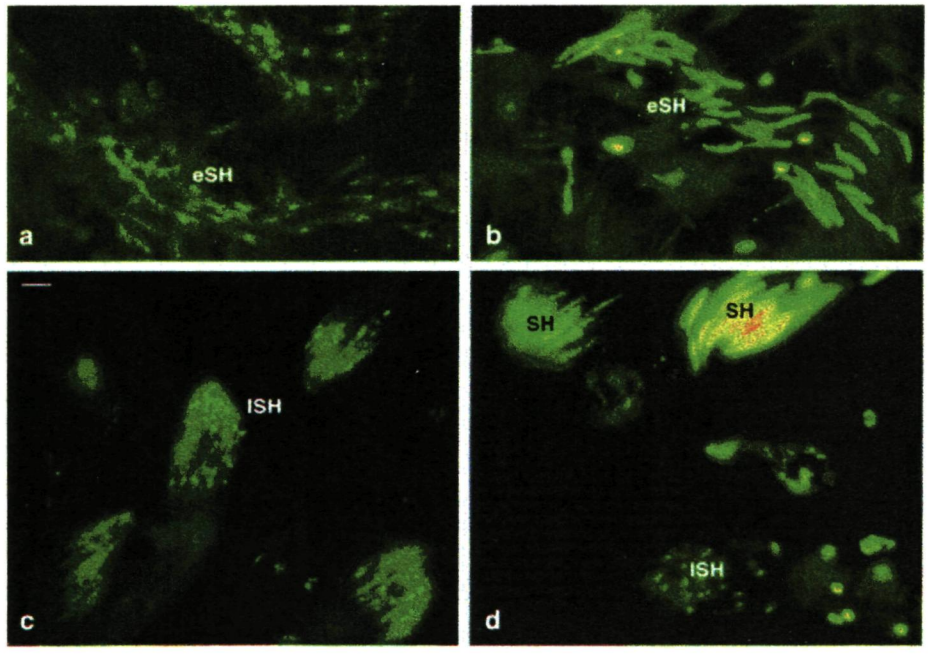

eSH
a
eSH
b
ISH
c
SH
SH
ISH
d

Fig 9. Early spermatids of *D hydei* after TCA treatment
a,b Pre-elongation spermatids after the incubation with the anti-H3 3p antibodies **a** Immunofl uorescence **b** Phase contrast The protein body shows strong reaction Some labeling in the chromatin on one side of the protein body is also observed
c,d Pre-elongation spermatids after incubation with the anti-H3 serum **c** Immunofluorescence **d** Phase contrast The chromatin, but not the protein body displays reaction with the serum
e, f Elongating spermatids after the incubation with the anti-H3 3p antibodies **e** Immunofluorescence **f** Phase contrast The reaction is observed in the protein body and in the chromatin, lying directly under the nuclear membrane
g, h Immunofluorescence of the elongating spermatids after the incubation with the anti-H3 antibodies The labeling is distributed throughout the nucleus The protein body shows almost no label **Pb** - protein body, **N** - nucleus, **M** nuclear membrane, **Nk** - nebenkern Bar represents 10 μm

Fig. 10. Spermatids of *D melanogaster* after TCA treatment
a, c Immunofluorescence after the incubation with the anti-H3 3p antibodies
b, d Immunofluorescence after the incubation with the anti-H3 antibodies **eSH** - elongating spermatid heads, **SH** - fully elongated spermatid heads in a bundle **lSH** - late (post-elongation) stage of spermatid head development Note the disapearance of the histone H3 during late stages of spermatid development Bar represents 10 μm

serum against total H3 (compare Fig. 10a and 10b). This, however, might be due to the fact that the i13.3-specific signal was much weaker.

Staining of the post-elongation spermatid cysts with both anti-H3 and anti-H3.3p antibodies showed that a transition occurs from almost even distribution of the label to a rather dispersed pattern (Fig. 10c and 10d). In some cysts label remains only in small patches shifted to the anterior part of the spermatid nuclei. This observation most likely reflects the substitution of the histones with other proteins. Such substitution of histones by arginine-rich proteins was observed by histochemical methods (Das et al., 1964; Hauschteck-Jungen and Hartl, 1982). It was shown to occur in post-elongation spermatids and to initiate at the posterior end of the nucleus (Hauschteck-Jungen and Hartl, 1982).

No staining was observed with the antibodies against core histones (H3, H3.3, H2A, H2B and H4) in the TCA-treated preparations of the mature sperm, prepared from the seminal vesicle (not shown). This indicates that the histones are either completely substituted by argine-rich proteins or inaccessible to antibodies. The first possibility seems more likely, since we could observe the gradual dissapearance of the histone H3-specific label in post-elongation spermatids.

4. Discussion

Histone replacement variants are present in most eukaryotes. Some of them, like histone H2A.F/Z and H3.3, are highly conserved in evolution. For the histone H2A.F/Z it is shown, that this variant is preferentially localized in active chromatin (Stargell et al., 1993) and that it is essential in *Tetrahymena* and in *Drosophila* (Liu et al., 1996; van Daal and Elgin, 1992). The function of the histone H3.3 variant is so far unclear. One way to address this function is to study the distribution of the H3.3 protein in chromatin. This can be achieved with the help of an antiserum, specific for the H3.3 variant.

4.1. Specificity of the anti-H3.3p antibodies

Only four amino acid substitutions distinguish the replacement histone H3.3 variant from the cell-cycle regulated variants H3.1 and H3.2. Three of these substitutions are clustered in the region 87-90 of the histone H3 amino acid sequence, making it an obvious place for targeting of antibodies which could specifically recognize the H3.3 variant but not its cell-cycle regulated counterparts (Fig. 1). It was demonstrated that even histone isoforms, differing only by the acetylation of particular lysines in the N-terminal domain, can be distinguished with the aid of polyclonal antisera raised against synthetic peptides, containing acetylated lysine residues (Turner and Fellows, 1989; Turner et al., 1989, Turner et al., 1992). This encouraged us to develop an H3.3-specific antiserum by immunization of rabbits with a synthetic peptide, covering the region 87-90.

When tested with histone peptides, the anti-H3.3p antibodies reacted equally

well with both the H3.1/2 and the H3.3 peptides. However, they became highly specific for the H3.3 peptide after preincubation with the H3.1/2 peptide (Fig. 2). The specificity of the antiserum was tested further on Western blots with calf thymus histone proteins, where the histone H3 variants were separated from each other in the AUT gel system (Fig. 3). The relative positions of all histone subtypes from the calf thymus mix were extensively characterized (Zweidler, 1984; Boulikas, 1985) and were used by us as a reference. The assignment of individual histone fractions was confirmed by using antisera, specific for each individual core histone. Our anti-H3.3p antibodies reacted mainly with one protein, which corresponds in mobility to the histone H3.3 variant. The reaction could be completely inhibited by the preincubation with the H3.3 peptide, but not with the H3.1/2 peptide. The same protein reacted, though weakly, with an antiserum, raised against the total histone H3 fraction (anti-H3 serum), as one could expect for a histone H3 subtype. The reaction of the anti-H3 antibodies with this protein was enhanced when these antibodies were purified over the H3.3 peptide, confirming the idea that it is indeed histone H3.3. It should be noted that these anti-H3 antibodies could not be made H3.3-specific by affinity purification over the H3.3 peptide, since even after such purification they still cross-reacted with histone H3.1. This is not surprising, since 5 out of 9 amino acids in the H3.3 peptide are also present in the H3.1 variant protein.

Taken together, our data indicate that the anti-H3.3p antibodies are specific for the H3.3 replacement variant. This is quite remarkable, because the antibodies bind equally well to both H3.1/2 and H3.3 peptides, yet on Western blots they react almost exclusively with the histone H3.3 but not with H3.1, although the latter is much more abundant in the calf thymus mix. The reason of such a difference in reactivity between the peptides and the whole histone H3 protein is unclear. Most likely, the conformation of the H3.1/2 peptide in solution is distinct from the conformation of the corresponding region of the H3.1/2 histone protein. This might explain our failure to produce an H3.1/2-specific antiserum.

4.2. SLE-derived antibodies with specificity to histone H3.3

SLE is an autoimmune disease which is characterized by the presence of a wide variety of auto-antibodies (for review see Monestier and Kotzin, 1992; Kramers et al., 1993; Tax et al., 1995). Among a number of different targets, antibodies against nuclear antigens like DNA, histones or nucleosomes are often found. It has been shown that anti-DNA antibodies are complexed to DNA and histones (or nucleosomes) and interact with heparan sulfate in the glomerular basement membrane (see Kramers et al., 1994 and references therein). This deposition of immunoglobulins and complexed molecules leads to nephritis, one of the most serious manifestations of the disease. The nucleosome seems to be the antigen driving the autoimmune response in SLE. Therefore the specificity of anti-nucleosome antibodies has been a subject of intensive study.

We have analysed the ability of three MAbs, derived from lupus prone mice, to distinguish between different histone H3 variants. Two of these MAbs, #56

and KM2, showed strong affinity to the H3.3, but not to the H3.1 histone variant on Western blot. In ELISA experiments, the MAb #56 displays strong reactivity with double stranded DNA, with the histone H3 peptide 83-100 (mostly with its dimer) and with mononucleosomes (Kramers, 1995; Stemmer et al., 1996). On Western blot this MAb reacts almost exclusively with the H3.3 histone. MAb KM2 strongly reacts with histone H3.3 on Western blot, but it also recognizes H2A and H4 histones. This is in agreement with the data from ELISA, showing that KM2 reacts with the N-terminal domains of histones H2A and H4 (van Bruggen et al., 1996).

The existense of auto-antibodies with affinity to the H3.3 variant in SLE mice is not surprising, since many mouse tissues contain the histone H3.3 protein (Zweidler, 1984). This histone variant has the interesting property of gradual accumulation in non-dividing tissues. With increasing age it can even become the prevalent histone H3 subtype in some of the tissues, for example, in liver or in cortical neurons (Zweidler, 1984, Pina and Suau, 1987). This might contribute to its capacity to serve as an autoantigen in adults, since its abundance in young individuals at the time when tolerance against autoantigens is established, is low. Analysis of a more representative group of SLE-derived antibodies is necessary to get insight into the role of the H3.3 variant as an autoantigen in SLE.

4.3. Application of the anti-H3.3p antibodies for immunocytology

4.3.1. Accessibility of the H3.3-specific epitope in chromatin

The current model of the nucleosome structure predicts that the amino acid residues 87-90, the target region of the H3.3 antibody, are on the surface of the histone octamer core, but covered by the DNA (see Fig. 4). In agreement with this model, our experiments have demonstrated, that the H3.3 specific epitope is inaccessible in the presence of DNA.

The necessity to remove the DNA prior to antibody binding creates obvious problems for the interpretation of the data. Hot TCA treatment, though rather unselective, is probably too harsh for decondensed chromatin, which might be completely dissolved and lost. The fact that we could detect with a serum against total histone H3 on TCA-treated testis squashes only the condensed part of chromatin supports this assumption. However, the TCA treatment is essential to obtain reproducible reactions of the anti-H3.3p antibodies and some other antibodies directed against the core histones with compacted chromatin. For example, no nuclear labeling of postmeiotic male germ cells was obtained with the histone H4 serum, used in this study, unless the tissues were treated with TCA. After the TCA treatment this anti-H4 antiserum gave a pattern identical to that obtained with the anti-H3 antiserum on untreated testis squashes (not shown). Apparently, when antisera are raised against purified core histones, many epitopes, recognized by these antisera, are concealed by the nucleosomal DNA. This observation also explains some previous failures in immunological detection of histones in condensed autosomes and X-chromosome of the *Droso-*

phila primary spermatocytes (Rungger-Brändle et al., 1981).

As an alternative way of DNA degradation, DNase I digestion was used. Results of this treatment have to be interpreted with even more caution, since the differential accessibility of active and inactive chromatin to this enzyme is a generally accepted fact. To overcome this problem at least partially, a prolonged digestion with a very high enzyme concentration was used.

4.3.2. Histone H3.3 distribution in polytene chromosomes

Anti-H3.3p antibodies gave no specific pattern on polytene chromosomes. Our data indicate that the H3.3 histone is only a minor component of these chromosomes with a rather uniform distribution. This argues against a specific function of the H3.3 variant in salivary gland nuclei.

Analyses of the distribution of another histone replacement variant, H2AvD, which is a *Drosophila* homologue of the the histone H2A.F/Z, indicated a widespread, but not general presence of this histone in the polytene chromosomes, with accumulation at particular genomic loci (van Daal and Elgin, 1992). Earlier studies with the antibodies against H2A.2 (D2), which probably corresponds to the H2AvD, showed its preferential accumulation in the interbands, which are supposed to be transcriptionally active (Donahue et al., 1986). The interpretation of the observed differences in distribution of the H2A and H3 replacement variants is not straightforward, because the experiments with anti-H2A antibodies were performed without removing the DNA, which might have influenced the accessibility of these proteins in certain chromosomal domains.

4.3.3. Histone H3.3 in *Drosophila* male germ line chromatin

Our previous studies have shown that one of the two *Drosophila* genes, encoding histone H3.3 (H3.3A), is strongly expressed in the testes (Akhmanova et al., 1995). Transcript in situ hybridization with the H3.3A-specific probes demonstrated that this gene is transcribed in spermatogonia and in primary spermatocytes (Akhmanova, unpublished).

To analyse the distribution of the histone H3.3 variant within male germ cell chromatin, we compared the reactions, obtained with the anti-H3.3p serum with those obtained with the an anti-H3 serum. The latter was raised against the total H3 protein and was expected to react with all histone H3 variants, because they share most of their epitopes.

We could not detect the H3.3 protein in spermatogonia despite of the fact that it is transcribed in these cells. Spermatogonia undergo quick mitotic divisions and, therefore, strongly express the cell-cycle regulated H3. Histone H3.3 represents probably only a minor part of the total histone H3 in these cells and the sensitivity of our experiments was insufficient to detect it.

Comparison of the reactions of the TCA-treated spermatocytes and young spermatids with the anti-H3 and anti-H3.3p sera indicates that in the chromatin of these cells the histone H3.3 distribution is not uniform. Only part of the chromatin, which reacts with the anti-H3 serum, reacts also with the H3.3 se-

rum. This is particularly apparent in primary spermatocytes, where most of the condensed chromatin does not react with the anti-H3.3p antibodies.

The strong transcription of the histone H3.3A gene in primary spermatocytes makes it an attractive hypothesis that histone H3.3 is incorporated in the active chromatin of these cells, substituting the cell-cycle regulated H3. The transcriptional activity in *Drosophila* spermatocytes is very high, since it has to provide the mRNAs for all subsequent steps of spermatogenesis, including the postmeiotic differentiation. Such intensive transcription might require or cause significant histone replacement. Histone replacement might also be required for transcriptional activation of the normally heterochromatic genes, such as the lampbrush loop forming genes of the Y-chromosome. Some support for this idea can be found in the literature. For example, biochemical studies have shown that newly synthesized nucleosomal H3.3 protein of chicken immature erythrocytes is preferentially found in active or competent chromatin fragments (Hendzel and Davie, 1990). The plant analog of the animal histone H3.3, the H3.2 or H3.III variant, also shows preferential accumulation in active chromatin (Waterborg, 1991, 1993).

To see, whether the correlation between transcriptional activity and the presence of H3.3 variant holds true in *Drosophila* spermatocytes, we have tried to remove the DNA enzymatically. Treatment with DNase I is milder than treatment with hot TCA, which destroys most of the intranuclear structures. We hoped that the DNase digestion will make the core histone epitopes available for reaction with antibodies, without fully destroying active chromatin.

Strong signals with the anti-H3.3p antibodies was obtained in the DNase-digested spermatocyte nuclei of both *D. melanogaster* and *D. hydei*. This is in agreement with the idea of the preferential accumulation of the histone H3.3 variant in active chromatin. It should be noted that considerable part of the label was associated with particular Y-chromosomal lampbrush loops (pseudonucleolus and clubs in *D. hydei*). Most of the nuclear signals were dependent on the DNase digestion, although a very weak reaction of the pseudonucleolus loop of *D. hydei* with the anti-H3.3p antibodies was also observed in nuclei, not treated with DNase (not shown). The significance of these antibody reactions is supported by the fact the same lampbrush loops are also stained with antibodies against histones H2B and H4 (Fig. 8) and histone H1 (Fig. 10d in Kremer et al., 1986). Labeling of the Y-chromosomal lampbrush loops with antibodies against core histones was observed before (Rungger-Brändle et al., 1981).

However, the interpretation of such labeling patterns requires extreme caution. The main components of the lampbrush loops are not the chromatin fibrils but extremely long transcripts and proteins (for review see Hennig, 1987, Hennig et al., 1989). Therefore it is doubtful that the loop chromatin on itself can be the source of the observed strong reactions. One cannot exclude that the histones are at least partially bound to transcripts, although such a possibility would seem rather exotic. Based on immunological studies, a variety of proteins were reported to be present in the Y-chromosomal lampbrush loops,

including not only RNA-binding proteins (Glätzer, 1984), but also laminin B2 (Wang et al., 1992) and tektin (Pisano et al., 1993), which are not even expected to enter the nucleus. Accumulation of this type of data raises the question of a possible artifactual nature of some of the observed immunoreactions. We would like to conclude, therefore, that although our data indicate the presense of histones, including the H3.3 variant, in the lampbrush loops, additional studies are necessary to prove it. Epitope tagging of histone proteins might provide a solution to this problem.

Strong staining with anti-H3.3p antibodies was observed in the protein bodies in the nuclei of early spermatids. Similar correlation between the presense in the lampbrush loops before the meiois and in the nuclear protein bodies in postmeiotic cells was observed for histone H1 (Kremer et al., 1986) and some RNA-associated proteins (Glätzer, 1984). Little is known about the function of these structures. It was shown, that they undergo a charateristic sequence of changes in structure and size, which occur in parallel with the changes in the degree of chromatin condensation (Lindsley and Tokuyasu, 1980; Grond, 1984, Hennig and Kremer, 1990). They contain basic proteins and are devoid of DNA (Grond, 1984; Kremer et al. 1986). It is possible that the protein bodies serve as a reservoir or a sink for chromosomal proteins during chromatin rearrangement.

4.4. Histone substitution during sperm maturation in *Drosophila*

The ways of packaging of the genome in sperm cells are highly variable between different organisms (for review, see Poccia, 1986). While in some species somatic histones are only partially substituted (for example, *Xenopus*; see Dimitrov et al., 1994, and references therein), in many other species they are replaced by small very basic proteins, called protamines (for review, see Hecht, 1989b). Almost nothing is known about the composition of the chromatin in spermatozoa of *Drosophila*. Histochemical observations indicate the appearance of arginine-rich proteins in late spermatid nuclei (Das et al., 1964; Hauschteck-Jungen and Hartl, 1982). Our data, demonstrating that histones are removed in late spermatids, are complementary to these observations. We did not observe labeling of the mature sperm heads with any of the antisera against core histones even after prolonged treatment with hot TCA. This probably indicates their absense in mature sperm, although epitope masking cannot be completely excluded.

Taken together, the data presented in this paper demonstrate that the replacement histone H3.3 can be distinguished immunologically from its cell-cycle regulated counterparts. The anti-H3.3p antibodies can be used to analyse the localization of this histone variant in the chromatin. The possibility to detect the histone H3.3 protein in cytological preparations will be very useful for subsequent genetic analysis of the function of this variant. Due to the high degree of evolutionary conservation of the histone H3.3 our serum can be applied to different animal systems.

Acknowledgements. We are very grateful to Lisette Sanders and Denise Vlooswijk for the help with immunocytology experiments. We thank Petra Bindels, Wenli Gu and Harry Harhangi for help in conducting the experiments and valuable discussions. Financial support was provided by the European Community Human Capital Programme (Networks), contract No. CHRX-CT94-0511.

Chapter VII

Identification and characterization of the *Drosophila* histone H4 replacement gene

Anna Akhmanova, Koos Miedema and Wolfgang Hennig

Department of Molecular and Developmental Genetics, Faculty of Sciences, Catholic University of Nijmegen, Toernooiveld, 6525 ED Nijmegen, The Netherlands.

This chapter is published in:
FEBS Lett. 388, 218-222 (1996)

Identification and characterization of the *Drosophila* histone H4 replacement gene

Anna Akhmanova, Koos Miedema, Wolfgang Hennig*

Department of Molecular and Developmental Genetics Faculty of Sciences Catholic University of Nymegen Toernooiveld 6525 ED Nymegen The Netherlands

Received 2 May 1996

Abstract **Replacement variant genes for different histones have been described in most higher eukaryotes. However, so far no such gene has been found for histone H4. We have isolated from both *Drosophila melanogaster* and *D. hydei* a novel histone H4 encoding gene, H4r, which displays all the properties of a histone replacement variant gene: it contains introns, generates polyadenylated mRNA, represents the predominant H4 transcript in non-dividing tissues and is present in the genome as a single copy. The encoded polypeptide is identical to the *Drosophila* cell-cycle regulated histone H4. The fact that it is a single copy gene makes it prone to genetic analysis and it might be a useful tool for studying nucleosome structure and function.**

Key words Histone H4 Histone replacement gene, Evolution, *Drosophila melanogaster*, *Drosophila hydei*

1. Introduction

Histones in higher eukaryotes are encoded by two types of genes Genes for replication-dependent histones are active predominantly in the S phase of the cell cycle, they are usually present in multiple copies, contain no introns and their transcripts are not polyadenylated [1] Replacement histone genes are usually single copy, contain introns and produce polyadenylated mRNAs [1] All histone replacement genes characterized so far encode proteins which are at least in a few positions different from their replication-dependent counterparts Although cell-cycle regulated genes, encoding all five histones (H1, H2A, H2B H3 and H4), have been isolated from many organisms, in no species has a full complement of replacement genes been found For example, replacement variants of histone H3 were described in different animal species, plants and protozoa [2] Since newly synthesized histones H3 and H4 form a complex before they are assembled into chromatin [3] one would expect to find histone H4 replacement genes as well Until now no such gene has been described

In this study, we report the identification and characterization of histone H4 replacement genes from *Drosophila melanogaster* and *D hydei* The implications of our finding for understanding of histone gene evolution are discussed

2. Materials and methods

2 1 Reverse transcriptase-polymerase chain reaction (RT-PCR)

Total RNA was isolated from flies of both *Drosophila* species by the guanidinium-thiocyanate method [4] and reverse transcribed in the presence of oligo(dT) with the aid of Moloney Murine Leukemia Virus (M-MLV) reverse transcriptase (Gibco BRL) PCR was carried out on the resulting cDNA with oligo(dT)-adapter primer with the sequence 5'-GCGGATCCGAATTCATCG(T)$_{17}$- and a degenerate histone H4 primer mh4 I with the sequence 5'-AAAGATCTAYGAR-GARACNCGNGGNGT (see Fig 1A) For the amplification of the 5' part of the *D melanogaster* cDNA 10 µg of total RNA were reverse transcribed in the presence of 0 5 pmol of the oligonucleotide m4-I with the sequence 5' TTTGATCAAGTATTTTGAATGTAGT-TAAG The resulting RT mix was diluted 1 50 and oligo(dC) tails were added with the terminal transferase (Boehringer Mannheim) in the presence of 1 mM dCTP PCR was performed with oligo(dG)-adapter primer with the sequence 5'-AGCTCAGAGCGGGCC-GCAAGCTT(G)$_{12}$ and primer m4-II 5'-CAAGATCTTGCGCT-TGGCGTGCTCG For the amplification of the 5' part of the *D hydei* cDNA, a pool of total cDNA was C-tailed and PCR was performed with oligo(dG)-adapter primer and primer h4-I with the sequence 5'-GAGAATTCTTGGCTGCTGCTTCAG PCR products were subcloned and several independent subclones were sequenced

2 2 Screening of the genomic library

A *D melanogaster* cDNA fragment, obtained by PCR with oligo(dT) and primer mh4-I was labeled with [α-^{32}P]dCTP by PCR with the same primers and used as a probe to screen the cosmid genomic library of *D melanogaster* Canton S strain in Lawrist4 [5] Two high-density filters with arrayed cosmid clones of this library were kindly provided by Dr J Hoheisel The filters were hybridised in 0 5 M Na-phosphate buffer (pH 7 2) 7% SDS, 1 mM EDTA and 1% BSA at 60°C and washed in 0 1 M Na-phosphate buffer 0 1% SDS at 25°C After exposure to the X-ray film the filters were stripped and rescreened with a *D melanogaster* histone H2B probe derived from the cluster of the cell-cycle regulated histone genes ([6] positions 677–968) Clones which hybridized to the histone H4, but not the H2B probe, were analysed further

2 3 DNA sequencing and sequence analysis

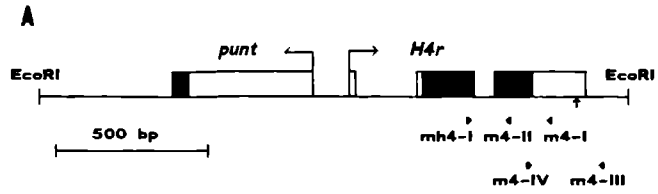
DNA sequence was determined on both strands with the help of the ALFexpress automatic sequencer (Pharmacia) Sequence analysis was performed using the programs of the GCG Sequence Analysis package [7] The number of synonymous substitutions per site (K_s) was calculated according to Nei and Gojobory [8] Standard errors (S E) were calculated according to Nei and Jin [9]

2 4 Northern blotting

Total RNA (10 µg per lane) was separated on 1 2% agarose-formaldehyde gel and transferred to Hybond nylon membrane (Amersham) Hybridization was performed under the same conditions as described for library screening and the filters were washed in 0 3 M Na phosphate buffer 0 1% SDS at 60°C All probes were labeled by PCR The *D melanogaster* H4r 3'UTR probe was amplified with primers m4-III 5'-GAGGATCCTAAGCGAGTTTGTTGGACCT and m4-IV 5'-TTACGGATTTGGCGGTTAAG (Fig 1A)

*Corresponding author Fax (31) (24) 3653390
E-mail kunmyl@caos kun nl

The nucleotide sequence data reported in this paper will appear in the EMBL GenBank and DDBJ nucleotide sequence databases with the following accession numbers *D melanogaster* mRNA for histone H4r — X97438 *D melanogaster* histone H4r gene — X97437 *D hydei* mRNA for histone H4r — X97436



B

```
          punt gene                <
-148 AAATCTTTTTTGTCGCAACGGGGCGGCGATAgcgatggccgataggcgataactgaccagaggctggtagtgagacaatag
                                                                  > H4r gene
 -68 aaatactagatacgcgcattttgcgaaaggtattttgaggtgtgcgagatctggtcacgctcacaagcAATTTTCTGGAC
  -3 ATTTCTTGAAgtgagttttgttgctgagaaatttaagtaaatatagaagaatcggtcaacaaagttttttgtgcataa
  93 acataaaaaaaataattggagaccattattttgtatgaaaacagacttattaacaacagttttcacctgtcaaatgaacg
 173 tttacctttcaccaaacatggcgtcgctttacggtttattgtgtcttctcttcttttcagCCAACTTTACGAAGCAACTG

       M  T  G  R  G  K  G  G  K  G  L  G  K  G  G  A  K  R  H  R  K  V  L  R  D  N   20
 253 AGAAATGACTGGTCGTGGAAAGGGTGGCAAAGGATTGGGAAAGGGGGGCGCCAAGCGTCATCGTAAGGTGCTTCGTGATA

       I  Q  G  I  T  K  P  A  I  R  R  L  A  R  R  G  G  V  K  R  I  S  G  L  I  Y    52
 333 ACATCCAGGGTATCACCAAGCCTGCTATTCGCCGTTTGGCTCGTCGCGGCGGTGTTAAGCGTATCTCTGGCTTGATTTAC

       E  E  T  R  G  V  L  K                                                          80
 413 GAGGAAACTCGCGGTGTGCTAAAGgtacgttctggcgaaacgaagggaaagtggtgaaaaactgattaatttctcggaaa

             V  F  L  E  N  V  I  R  D  A  V  T  Y  T  E  H  A  K  R  K  T  V  T  A    84
 493 attcacagGTATTCCTTGAGAACGTTATCCGTGACGCTGTCACCTACACCGAGCACGCCAAGCGCAAGACCGTGACCGCC

       M  D  V  V  Y  A  L  K  R  Q  G  R  T  L  Y  G  F  G  G  *                    103
 573 ATGGACGTGGTCTATGCCCTCAAGCGCCAGGGACGCACCCTTTACGGATTTGGCGGTTAAGCACGCCCTCTAACCTCCAC
 653 TTGATAAATAACCATGGAGTTAACTTAACTACATTCAAAATACTTGATCCAATTGAATCTTAATCAGTAGTATTCAACGC
 733 AAACTTAACCCAACGCAAATATTTTATAATTTCACAATATACGGACTAAATAAATTCAAAAGATATCACATAGAGAACTCA
 813 ATATCATGCATATGTGCAttttcacttggggcggttaatcatccgaaaataaggtccaacaaactcgcttagggagcaca
```

Fig 1 (A) Structure of the *D melanogaster* H4r gene Transcription start sites are indicated by horizontal arrows Black boxes represent translated parts and open boxes represent untranslated parts of the exons of the H4r and the *punt* gene The vertical arrow indicates the polyadenylation signal in the H4r gene Horizontal arrowheads indicate the positions of the oligonucleotides used for PCR (B) Nucleotide sequence and the encoded amino acid sequence of the *D melanogaster* histone H4r gene The nucleotides of the *D melanogaster* sequence are numbered at the left starting from the transcription initiation site, which is indicated in bold and italic Amino acids are numbered at the right and the stop codon is indicated by an asterisk Introns and the intergenic region between the H4r and the *punt* genes are shown in lower case letters The polyadenylation signal is underlined and the experimentally determined polyadenylation sites are shown in bold

3. Results and discussion

3 1 Isolation of the H4r cDNA and genomic clones

We have undertaken a systematic search for histone H4 replacement variants in *Drosophila* It was assumed that, due to the extremely high evolutionary conservation of histone H4, a replacement gene would encode a protein quite similar to cell-cycle regulated H4 and that its mRNA would be polyadenylated A degenerate oligonucleotide primer was designed from the histone H4 sequence and used in combination with oligo(dT) for RT-PCR on total RNA from flies One major fragment was obtained The 5′ part of the fragment encoded a polypeptide, identical to the C-terminus of the *Drosophila* histone H4 The 3′ part of the fragment was novel in its sequence and had no similarity to the 3′UTR of the cell-cycle regulated histone H4 It contained no inverted repeat sequence, characteristic for the replication-dependent histones Rather, it contained a polyadenylation signal near its 3′ end, as was expected for a replacement histone gene (Fig 1) We called this gene histone H4r (H4 replacement gene) The sequence of this fragment was used to design two new primers to isolate the 5′ part of the cDNA using the rapid amplification of cDNA ends procedure [10] One PCR product was obtained It was cloned and several independent subclones were sequenced The fragment overlapped with the 3′ cDNA fragment and was identical to it in the overlapping region The 5′ ends of the different subclones were the same, so we concluded that the isolated cDNA was complete in its 5′ part The same approach was used to isolate the H4r cDNA from *D hydei* The 5′UTR of this cDNA is probably not complete, because we used a total pool of oligo(dT)-primed *D hydei* cDNA to isolate it

The obtained H4r cDNA fragment was used to screen a cosmid genomic DNA library of *D melanogaster* The filters with arrays of cosmid clones, representing together four genomic equivalents, were hybridized to the probe under non-stringent conditions Approximately 60 clones gave a positive signal Most of these signals were the result of a cross-reaction of our probe with the histone H4 gene of the histone repeat cluster (which is present in the genome in ~100 copies [11]) We assumed that histone H4r gene is localised outside of the histone cluster, similar to the replacement genes

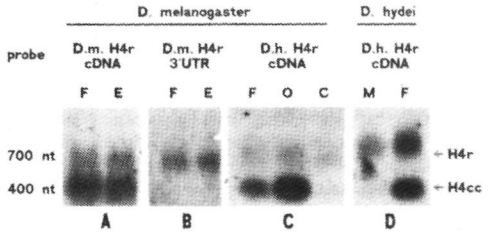


Fig. 2. Northern blot analysis of the H4r expression. Total RNA (10 μg) was prepared from: F — female flies, M — male flies, E — embryos, C — fly carcasses, O — ovaries. Length of the transcripts in nucleotides (nt) is indicated on the left.

H2AvD [12] and H3.3A and H3.3B [13]. To eliminate the cosmids containing the histone cluster copies, we rescreened the library with the cell-cycle regulated histone H2B probe and analysed those clones which did not react with it. Three of these clones, JHD1:72G2, JHD4:19G11 and JHD5:58C4, hybridized strongly to the probe. They proved to be overlapping cosmids, which contained the same 2 kb *Eco*RI fragment encompassing the entire H4r gene (Fig. 1). We also analysed several weaker hybridizing clones. In none of them was an H4-like gene found.

Taken together, the results of the PCR experiments on the cDNA pool and the genomic library screen suggest that the isolated gene, H4r, is the only replacement histone H4 gene in the *D. melanogaster* genome. Also genomic Southern blotting (results not shown) and Northern blotting (see below) provided no indication for the presence of additional copies of the H4-encoding replacement histone gene in this species. The fact that the H4r is a single copy gene makes it susceptible to genetic analysis. It provides a good opportunity to perform genetic experiments with a histone H4 in a higher eukaryote.

3.2. *Genomic organization of the* D. melanogaster *histone H4r*

Comparison of the cDNA and genomic sequence of the histone H4r showed that the gene contains two introns; one of ~200 bp interrupts the 5'UTR while the second one is located in the ORF of the gene (Fig. 1). This structure is quite similar to that of the *Drosophila* H3.3B genes except that the first intron of H4r is much shorter than that of histone H3.3B [13]. The small size of the H4r 3'UTR makes this gene more similar to the H3.3A than to the H3.3B gene.

Analysis of the sequence flanking the histone H4r gene showed that there is another gene, *punt*, which is localized immediately upstream from the H4r gene and is transcribed from the opposite DNA strand (Fig. 1A). *punt* encodes a type II *dpp* receptor and is expressed at all developmental stages [14,15]. The transcription start sites of the two genes are separated by only ~120 bp. This does not leave much space for the promoters of the two genes and suggests that they might be to a certain extent coregulated. Similarly close linkage to a housekeeping gene, encoding the oligosaccharyltransferase 48 kDa subunit, was also observed for the *D. melanogaster* histone H3.3B [13].

Analysis of the promoter region revealed no presence of a TATA-box around position −25, which is again similar to the promoter of H3.3B gene. The GAGA-factor binding sites [16], present in the *Drosophila* cell-cycle regulated H4 promoter, were not found either. Further, we could not identify any promoter elements, common for the *Drosophila* H2AvD, H3.3B and H4r genes. So even if these histone replacement genes are coregulated, it is not apparent from the comparison of their promoter sequences.

The H4r gene was mapped by hybridizing the 2 kb *Eco*RI genomic fragment to the polytene chromosomes (not shown) and was found in the locus 88C of chromosome 3. This is in agreement with the localization of the *punt* gene [14,15].

3.3. *Analysis of the histone H4r gene expression by Northern blotting*

Hybridization of the *D. melanogaster* H4r cDNA probe to Northern blots with total fly and embryonic RNA revealed two bands of ~400 nt and ~700 nt (Fig. 2A). The ~400 nt band corresponds to the cell-cycle regulated histone H4 (H4cc) mRNA, while the ~700 nt band corresponds to the H4r transcript. This was confirmed by probing the same blot with the H4r 3'UTR probe (Fig. 2B). The cell-cycle regulated H4 mRNA is more abundant both in embryos and in females. The latter is caused solely by the storage of the histone cell-cycle regulated messengers in the oocytes. To show that, we hybridized a Northern blot with *D. melanogaster* RNA from ovaries and fly carcasses (remainder of the fly after removing the ovaries) to the *D. hydei* H4r cDNA. This cDNA is almost equally diverged (~17%) from *D. melanogaster* H4 genes of both types in the ORF region and has no similarity to either of the genes in the untranslated regions. As one can see from Fig. 2C, H4cc mRNA is very abundant in the ovary, while the H4r is the predominant transcript in the fly carcass (somatic, mainly non-dividing tissues). This result is exactly as expected for a replacement histone and is very similar to that obtained for the H3.3 genes [13]. A single H4r transcript of ~800 bp was observed in *D. hydei* (Fig. 2D). This size is in agreement with a longer 3'UTR of the H4r cDNA in this species.

3.4. *Sequence comparison of the* Drosophila *histone H4 genes*

The *D. melanogaster* and *D. hydei* H4r cDNAs encode a protein, identical to the *Drosophila* cell-cycle regulated H4. This is the first case when a replication-dependent and a replacement histone gene of a particular species encode exactly the same polypeptide. The function of the replacement-type histone genes is not yet clear, but one can imagine two different possibilities. On the one hand, incorporation of the replacement variant histone proteins might be necessary to alter the structure of certain chromatin domains in the absence of replication. If this is true, one would expect to find functionally important differences between replacement and cell-cycle regulated histone proteins. On the other hand, it is possible that low-level synthesis of histones per se is necessary for chromatin maintenance in non-dividing cells. In the latter case the proteins encoded by the genes of the two types might be identical or differ by neutral amino acid substitutions. Our data on replacement gene of H4 support the second hypothesis. However, one should bear in mind that histones H3 and H4 are assembled into the chromatin as an $(H3–H4)_2$ tetramer. It is possible, therefore, that in order to incorporate into chromatin the replacement variant histone H3.3, replication-independent expression of histone H4 is required and that the H4r gene fulfils this function.

On the nucleotide level the four *Drosophila* H4 ORF sequences are quite divergent, because of the presence of multi-

ple synonymous substitutions As expected, phylogenetic analysis by different methods [7] indicates that the cell-cycle regulated and replacement genes of two species are more related to each other than the two types H4 genes within one species The number of synonymous substitutions per site between the H4r genes ($K_s = 1\ 30$ S E $= 0\ 16$) is somewhat higher than that between the cell-cycle regulated H4 genes ($K_s = 0\ 94$, S E $= 0\ 10$) This is in contrast to the values obtained for the *Drosophila* H3 histones, whose variants evolve slower than replication-dependent genes [13] The base composition at synonymous positions and the codon usage bias are quite similar between the two types of H4 genes This indicates that the type of genomic organization (single copy versus tandem repeats) and the mode of regulation do not have much influence on the evolution of these histone-encoding sequences in *Drosophila*

The untranslated regions of the *D melanogaster* and *D hydei* cDNAs display very little similarity The same lack of conservation was observed for the *Drosophila* H3 3 genes [13] This indicates that the evolutionary constraints on these regions of histone replacement genes are quite low in *Drosophila*, in contrast to the vertebrate genes [17,18]

3 5 Evolution of the H4 replacement gene

Replacement variants of core histones are thought to be evolutionarily ancient The histone H2A variant probably diverged from its cell-cycle regulated counterpart before the separation of fungi and ciliates from the rest of the eukaryotic phyla [19] The histone H3 3 variant probably arose very early in animal evolution, before the split between invertebrates and vertebrates [2,13] One would expect a similar evolutionary history for a histone H4 replacement gene

To establish the evolutionary relationships between genes, comparison of the intron positions is very useful Histone H4 genes were sequenced in many different species, but most of them do not contain introns The exceptions are the H4 genes of certain fungi (from *Neurospora*, *Aspergillus*, *Physarum*) The intron positions in all these organisms are different from that of the *Drosophila* H4r This correlates with their very early separation from the rest of the eukaryotes Recently, two nematode intron-containing H4 sequences were deposited in the data bases One of them, from *C elegans* cosmid C50F4 which was sequenced as a part of the *C elegans* Genome Sequencing Project [20], contains an intron at a position different from the H4r gene The other, an incomplete genomic sequence from *Ascaris lumbricoides* (EMBL accession numbers Z69289 and Z69290, P Duda et al , unpublished) contains an intron at a position identical to that of the *Drosophila* H4r The H4 gene from *Ascaris* displays other properties of a replication-independent histone gene — the absence of the 3′ terminal palindrome and the presence of the polyadenylation signal Assuming that this gene is expressed, one can conclude that it is the nematode homologue of the *Drosophila* H4r gene This indicates that the histone H4 replacement and cell-cycle regulated genes separated from each other early in animal evolution It is noteworthy that the polypeptide, encoded by the *Ascaris* gene is at four positions different from the *Drosophila* sequence while the two *Drosophila* histone H4 sequences are identical However, judging from the intron position and the 3′ regulatory sequences the replacement H4 genes from *Ascaris* and *Drosophila* are evolutionarily closer to each other than the replication dependent and the replacement H4 genes of *Drosophila* This would mean that histone H4 genes undergo coevolution within a species and that establishing the phylogenetic relationships between core histone genes based on protein alignments may be difficult This conclusion is strengthened by the fact that the intron containing H4 gene from the *C elegans* cosmid C50F4 also encodes a protein identical to the cell cycle regulated H4 of this species The fact that its intron is positioned differently from that in *Drosophila* and *Ascaris* H4 genes complicates the picture of the evolution of the replacement histone H4 Characterization of the replacement histone H4 genes from other higher eukaryotes will hopefully clarify this picture

Acknowledgements We are very grateful to Dr J Hoheisel for providing the cosmid library filters and the clones to Denise Vroomjk for her contribution to the analysis of the genomic clones and to Dr G Korge for help with the polytene chromosome hybridization We thank Petra Bindels Wenli Gu and Harry Harhangi for help in conducting the experiments and for valuable discussions

References

[1] Schumperli D (1986) Cell 45 471 472
[2] Thatcher T H MacGaffey J Bowen J Horowit S Shapiro D L and Gorovsky M A (1994) Nucleic Acids Res 22 180 186
[3] Perry C A Dadd C A Allis C D and Annunziato A T (1993) Biochemistry 32 13605 13614
[4] Chirgwin J M Przybyla A E MacDonald R J and Rutter W J (1979) Biochemistry 18 5294 5299
[5] Hoheisel J D Lennon G G Zehetner G and Lehrach H (1991) J Mol Biol 220 903 914
[6] Matsuo Y and Yamazaki I (1989) Nucleic Acids Res 17 225 238
[7] Devereux J Haeberli P and Smithies O (1984) Nucleic Acids Res 12 387 395
[8] Nei M and Gojobory T (1986) Mol Biol Evol 3 418 426
[9] Nei M and Jin L (1989) Mol Biol Evol 6 290 300
[10] Frohman M A (1990) in PCR Protocols a Guide to Methods and Applications (Innis M A Gelfand D H Sninsky J J and White T J eds) Academic Press San Diego pp 28 38
[11] Lifton R P Goldberg M L Karp R W and Hogness D S (1977) Cold Spring Harbor Symp Quant Biol 42 1047 1051
[12] van Daal A White E M Gorovsky M A and Elgin S C (1988) Nucleic Acids Res 16 7487 7497
[13] Akhmanova A S Bindels P C J Xu J Miedema K Kremer H and Hennig W (1995) Genome 38 596 600
[14] Childs S R Wrana J L Arora K Attisano L O'Connor M B and Massague J (1993) Proc Natl Acad Sci USA 90 9475 9479
[15] Ruberte E Marty T Nellen D Affolter M and Basler K (1995) Cell 80 889 897
[16] Gilmour D S Thomas G H and Elgin S C R (1989) Science 245 1487 1490
[17] Wells D Hoffman D and Kedes L (1987) Nucleic Acids Res 15 2871 2889
[18] Chalmers M and Wells D (1990) Nucleic Acids Res 18 3075
[19] van Daal A White E M Elgin S C R and Gorovsky M A (1990) J Mol Evol 30 449 455
[20] Wilson R Ainscough R Anderson K Baynes C Bonfield J Barton J Connel M Copsey T Cooper J et al (1994) Nature 368 32 38

Chapter VIII

Drosophila melanogaster histone H2B retropseudogene is inserted into a region rich in transposable elements

Anna Akhmanova and Wolfgang Hennig

Department of Molecular and Developmental Genetics, Faculty of Sciences, Catholic University of Nijmegen, Toernooiveld, 6525 ED Nijmegen, The Netherlands

Abstract

We have isolated and characterized the genomic sequence of a *Drosophila melanogaster* histone H2B gene, which is localized outside of the cluster of the replication-dependent histone genes. This gene has all the properties of a retropseudogene. It is highly homologous to the transcribed part of *D. melanogaster* histone H2B gene, but not to its flanking regions, and it is surrounded by short direct repeats. It also contains several point mutations which preclude its translation. The sequence of the 3' part of this pseudogene is compatible with the possibility that the 3' terminal stem-loop structure has served as a primer for the reverse transcription event, from which this pseudogene has originated. Analysis of the regions, flanking the histone H2B pseudogene, revealed the presence of three different types of transposable elements, suggesting that this genomic locus represents a hotspot for transposition.

1. Introduction

Replication-dependent histones of higher eukaryotes are encoded by multicopy genes. In birds and mammals these genes are arranged in random clusters, interspersed by relatively large regions of unrelated sequences (for recent papers, see Tönjes et al., 1989; Gruber et al., 1990; Albig et al., 1993; Drabent et al., 1993; Doenecke et al., 1994; Dong et al., 1994; Brown et al., 1996). In *Drosophila*, as well as in sea urchin (Kedes, 1979) and newt (Stephenson et al., 1981), histone genes are organized as tandemly repeated clusters. Each cluster contains all genes for the core histones and for H1. In *D. melanogaster* there are about 100 copies of cluster repeats, which are localized in region 39 DE on chromosome 2 (Lifton et al., 1978). There are two types of repeat units, which differ only by a short insertion in the H1-H3 spacer (Matsuo and Yamazaki, 1989).

In addition to the histone genes, present in the cluster, single copy histone genes are also present in the *D. melanogaster* genome. These include the genes, encoding replacement histone variants *H2AvD* (van Daal et al., 1988), *H3.3A* and *H3.3B* (Akhmanova et al., 1995) and *H4r* (Akhmanova et al., 1996). Early hybridization studies suggested that isolated copies of histone H2B might also be present in *Drosophila* genome (Childs et al., 1981). We have searched for such an "orphon" H2B gene by looking for genomic clones, that would hybridize to the H2B, but not to the other histone probes. One such genomic sequence was isolated. It contains a histone H2B retropseudogene of an apparently recent origin, which is surrounded by transposable elements of different types.

2. Materials and Methods

2.1. Screening of the genomic library

A cosmid genomic library of *D. melanogaster* Canton S strain in Lawrist4 vector (Hoheisel et al., 1991) was used for screening. Two high-density filters with arrayed cosmid clones of this library were kindy provided by Dr. J. Hoheisel. The histone H2B probe corresponded to the positions 677-968 of the cluster of the replication-dependent histones (Matsuo and Yamazaki, 1989). The screening

procedure was described before (Akhmanova et al., 1996). The cosmid clones which hybridized to the histone H2B but not to the histone H4 probes were analysed by Southern blotting.

2.2. DNA sequencing and sequence analysis

DNA sequence was determined on both strands with the help of the ALFexpress automatic sequencer (Pharmacia). Sequence analysis was performed using the programs of the GCG sequence analysis package. The number of synonymous and non-synonymous substitutions per site was calculated according to Nei and Gojobory, 1986.

3. Results and Discussion

3.1. An "orphon" histone H2B sequence is a retropseudogene

Clones from a genomic cosmid library, representing about four genomic equivalents of *Drosophila*, were successively screened with a histone H4 and a histone H2B probe (Akhmanova et al., 1996). Three clones hybridized only to the histone H2B probe and, therefore, were unlikely to originate from the histone repeat cluster, where all five histone genes are tightly linked. These three clones, JHD1:71B4, JHD5:57H7 and JHD5:64A1 were analysed in detail. They proved to be overlapping cosmids, containing the same 3 kb HincII fragment which hybridized to the histone H2B probe. This fragment was sequenced (see Fig. 1 and 2).

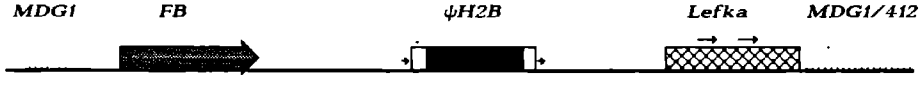
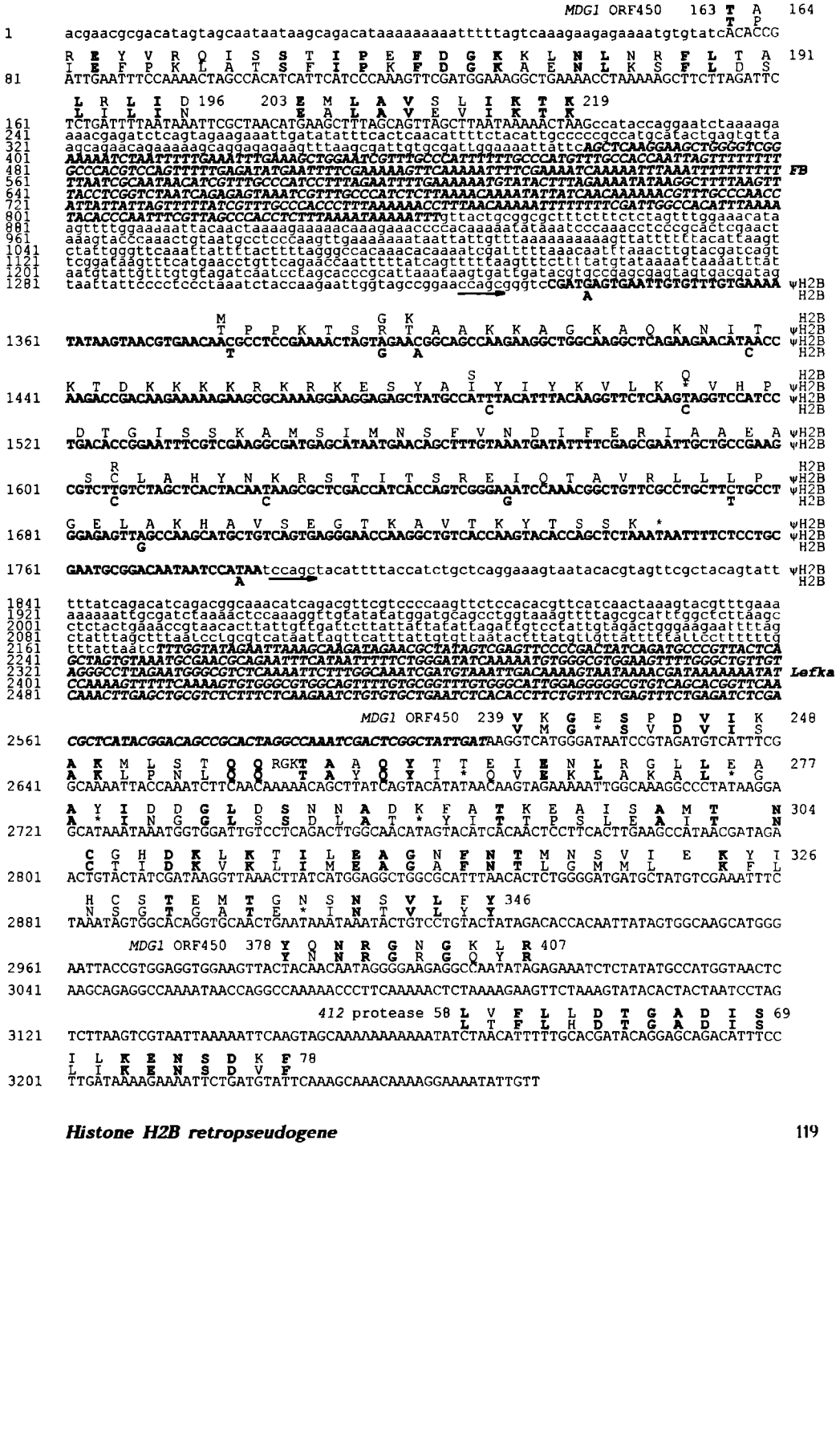


Fig. 1. Structure of the genomic DNA fragment, containing the *ψH2B* gene. Black box represents the region, homologous to the open reading frame of the H2B gene and open boxes represent sequences, homologous to the untranslated regions of the H2B gene. Small arrows indicate the positions of direct repeats, surrounding the *ψH2B* gene. Big grey arrow indicates the position of the *FB* transposon. Cross-hatched box represents the *Lefka*-related sequence, with a direct repeat in it indicated by horizontal arrows. Stippled boxes represent regions similar to *MDG1* and *412* retrotransposons.

Fig. 2. Sequence of the *ψH2B* gene and flanking regions. Sequences with similarity to other *D melanogaster* sequences are shown in upper case, and the rest in lower case. Nucleotides are numbered at the left. *ψH2B* sequence is shown in bold, with the encoded amino acids within the H2B ORF indicated above the sequence. Asterisks represent stop codons. The nucleotides, which are different in the functional H2B gene, are shown below the sequence, and in case of non-synonymous substitutions, the differing amino acids are shown above the sequence. Direct repeats, flanking the *ψH2B* gene, are indicated by horizontal arrows. Foldback (*FB*) and *Lefka* sequences are shown in bold and italics. For *MDG1* and *412* retrotransposon-related sequences, the encoded amino acids are shown in the regions of significant similarty (determined with a BLASTX search, GCG sequence analysis package). The corresponding parts of the retrotransposon ORFs are shown above the sequence and numbered at the right. Identical amino acids are shown in bold.


MDG1 ORF450 163 T A 164
T P
1 acgaacgcgacatagtagcaataataagcagacataaaaaaaaattttagtcaaagaagagaaaatgtgtatcACACCG
R E Y V R Q I S S T I P E F D G K K L N L N R F L T A 191
I E F P K L A T S F I P K F D G K A E N L K S F L D S
81 ATTGAATTTCCAAAACTAGCCACATCATTCATCCCAAAGTTCGATGGAAAGGCTGAAAACCTAAAAAGCTTCTTAGATTC
L R L I D 196 203 E M L A V S L I K T K 219
L I L I N E A L A V E V I K T K
161 TCTGATTTTAATAAATTCGCTAACATGAAGCTTTAGCAGTTAGCTTAATAAAAACTAAGccataccaggaatctaaaaga
241 aaacgagatctcagtagaagaaattgatatatttcactcaacatttttctacattgcccccgccatgcatactgaqtgtta
321 agcagaacagaaaaagcaggagagaagtttaagcgattgtgcgattggaaaattattcAGCTCAAGGAAGCTGGGGTCGG
401 AAAAATCTAATTTTTGAAATTTGAAAGCTGGAATCGTTTGCCCATTTTTTGCCCATGTTTGCCACCAATTAGTTTTTTTT
481 GCCCACGTCCAGTTTTTGAGATATGAATTTTCGAAAAAGTTCAAAAATTTTCGAAAATCAAAAATTTAAATTTTTTTTTT FB
561 TTAATCGCAATAACATCGTTTGCCCATCCTTTAGAATTTTGAAAAAATGTATACTTTAGAAAAATATAAGGCTTTTAAGTT
641 TACCTCGGTCTAATCAGAGAGTAAATCGTTTGCCCATCTCTTAAAACAAAATATTATCAACAAAAAACGTTTGCCCAACC
721 ATTATTATTAGTTTTTATCGTTTGCCCACCCTTTAAAAAACCTTTAACAAAAATTTTTTTTCGATTGGCCACATTTAAAA
801 TACACCCAATTTCGTTAGCCCACCTCTTTAAAATAAAAATTTgttactgcggcgctttctttctctagtttggaaacata
881 agttttggaaaaattacaactaaaagaaaaacaaagaaacccacaaaaatataaatcccaaacctcccgcactcgaact
961 aaagtacccaaactgtaatgcctcccaagttgaaaaaaatattattgtttaaaaaaaaaagttatttttttacattaagt
1041 ctattgggttcaaattattttacttttagggccacaaacacaaaatcgatttttaaacaatttaaacttgtacgatcagt
1121 tcggataagtttcatgaacctgttcagaaccaatttttatcagttttttaagtttcttttatgtataaaattaaaatttat
1201 aatgtattgttttgtgagatcaatcctagcacccgcattaaataagtgattgatacgtgccgagcgagtagtgacgatag
1281 taattattcccctcccctaaatctaccaagaattggtagccggaaccagcgggtcCGATGAGTGAATTGTGTTTGTGAAAA ψH2B
A H2B
M G K H2B
T P P K T S R T A A K K A G K A Q K N I T ψH2B
1361 TATAAGTAACGTGAACAACGCCTCCGAAAACTAGTAGAACGGCAGCCAAGAAGGCTGGCAAGGCTCAGAAGAACATAACC ψH2B
T G A C H2B
S Q H2B
K T D K K K K R K R K E S Y A I Y I Y K V L K * V H P ψH2B
1441 AAGACCGACAAGAAAAAGAAGCGCAAAAGGAAGGAGAGCTATGCCATTTACATTTACAAGGTTCTCAAGTAGGTCCATCC ψH2B
C C H2B
D T G I S S K A M S I M N S F V N D I F E R I A A E A ψH2B
1521 TGACACCGGAATTTCGTCGAAGGCGATGAGCATAATGAACAGCTTTGTAAATGATATTTTCGAGCGAATTGCTGCCGAAG ψH2B
R H2B
S C L A H Y N K R S T I T S R E I Q T A V R L L L P ψH2B
1601 CGTCTTGTCTAGCTCACTACAATAAGCGCTCGACCATCACCAGTCGGGAAATCCAAACGGCTGTTCGCCTGCTTCTGCCT ψH2B
C C G T H2B
G E L A K H A V S E G T K A V T K Y T S S K * ψH2B
1681 GGAGAGTTAGCCAAGCATGCTGTCAGTGAGGGAACCAAGGCTGTCACCAAGTACACCAGCTCTAAATAATTTTCTCCTGC ψH2B
G H2B
1761 GAATGCGGACAATAATCCATAAtccagctacattttaccatctgctcaggaaagtaatacacgtagttcgctacagtatt ψH2B
A H2B
1841 tttatcagacatcagacggcaaacatcagacgttcgtccccaagttctccacacgttcatcaactaaagtacgtttgaaa
1921 aaaaaattgcgatctaaaactccaaaggtttgtatatatggatgcagcctggtaaagttttagcgcatttggctcttaagc
2001 ctctactgaaaccgtaacactattgttgattcttattattatattagattgtcctattgtagactgggaagaattttag
2081 ctatttagctttaatcctgcgtcataattagttcatttattgtgttaatactttatgttgttatttttattcctttttg
2161 tttattaatcTTTGGTATAGAATTAAAGCAAGATAGAACGCTATAGTCGAGTTCCCCGACTATCAGATGCCCGTTACTCA
2241 GCTAGTGTAAATGCGAACGCAGAATTTCATAATTTTTCTGGGATATCAAAAATGTGGGCGTGGAAGTTTTGGGCTGTTGT
2321 AGGGCCTTAGAATGGGCGTCTCAAAATTCTTTGGCAAATCGATGTAAATTGACAAAAGTAATAAAACGATAAAAAAATAT Lefka
2401 CCAAAAGTTTTTCAAAAGTGTGGGCGTGGCAGTTTTGTGCGGTTTGTGGGCATTGGAGGGGGCGTGTCAGCACGGTTCAA
2481 CAAACTTGAGCTGCGTCTCTTTCTCAAGAATCTGTGTGCTGAATCTCACACCTTCTGTTTCTGAGTTTCTGAGATCTCGA
MDG1 ORF450 239 V K G E S P D V I K 248
V M G * S V D V I S
2561 CGCTCATACGGACAGCCGCACTAGGCCAAATCGACTCGGCTATTGATAAGGTCATGGGATAATCCGTAGATGTCATTTCG
A K M L S T Q Q RGKT A A Q Y T T E I E N L R G L L E A 277
A K L P N L Q Q T A Y Q Y I * Q V E K L A K A L * G
2641 GCAAAATTACCAAATCTTCAACAAAAACAGCTTATCAGTACATATAACAAGTAGAAAAATTGGCAAAGGCCCTATAAGGA
A Y I D D G L D S N N A D K F A T K E A I S A M T N 304
A * I N G G L S S D L A T * Y I T T P S L E A I T N
2721 GCATAAATAAATGGTGGATTGTCCTCAGACTTGGCAACATAGTACATCACAACTCCTTCACTTGAAGCCATAACGATAGA
C G H D K L K T I L E A G N F N T M N S V I E K Y I 326
C T I D K V K L I M E A G A F N T L G M M L K F L
2801 ACTGTACTATCGATAAGGTTAAACTTATCATGGAGGCTGGCGCATTTAACACTCTGGGGATGATGCTATGTCGAAATTTC
H C S T E M T G N S N S V L F Y 346
N S G T G A T E * I N T V L Y Y
2881 TAAATAGTGGCACAGGTGCAACTGAATAAATAAATACTGTCCTGTACTATAGACACCACAATTATAGTGGCAAGCATGGG
MDG1 ORF450 378 Y Q N R G N G K L R 407
Y N N R G R G Q Y R
2961 AATTACCGTGGAGGTGGAAGTTACTACAACAATAGGGGAAGAGGCCAATATAGAGAAATCTCTATATGCCATGGTAACTC
3041 AAGCAGAGGCCAAAATAACCAGGCCAAAAACCCTTCAAAAACTCTAAAAGAAGTTCTAAAGTATACACTACTAATCCTAG
412 protease 58 L V F L L D T G A D I S 69
L T F L H D T G A D I S
3121 TCTTAAGTCGTAATTAAAAATTCAAGTAGCAAAAAAAAAAAATATCTAACATTTTTGCACGATACAGGAGCAGACATTTCC
I L K E N S D K F 78
L I K E N S D V F
3201 TTGATAAAAGAAAATTCTGATGTATTCAAAGCAAACAAAAGGAAAATATTGTT

Sequence analysis has shown that the 3 kb HincII fragment contained a region approximately 450 nt long which is 97% similar to the *D. melanogaster* histone H2B gene. This H2B gene is unlikely to be functional because of a point mutation in the translation start codon (ATG →ACG) and a nonsense mutation in the Gln 44 codon (CAG →TAG) (Fig. 2). We can conclude, therefore, that the isolated histone H2B-homologous sequence is a pseudogene, which we will call *ψH2B*.

The homology between *ψH2B* and the normal histone H2B gene is confined to a region from position 371 to position 820 of the histone cluster (Matsuo and Yamazaki, 1989). Although the histone H2B transcription unit is not precizely mapped, there are good reasons to believe that this region is within the limits of the normal H2B mRNA. First, the 3' end of the H2B mRNA at the position 840 is determined by the highly conserved stem-loop sequence (Marzluff, 1992). Second, the localization of the transcription start for the histone H2B gene is likely to be in the region 365-370 of the histone cluster. Such a position for transcription start is in agreement with the length of the H2B transcripts determined by Northern blotting (Akhmanova, unpublished). Also, there is a TATA-box about 25 bp upstream of this region, which is the common distance between this element and the transcription start in eukaryotic promoters (Breathnach and Chambon, 1981). Therefore, the region of similarity between *ψH2B* and its functional counterpart is confined to the transcribed part of the histone H2B gene. It seems likely that *ψH2B* is a retropseudogene, which was derived from histone H2B mRNA by reverse transcription. This idea is supported by the presence of short (5 bp) direct repeats at the 5' and 3' end of *ψH2B* (see Fig. 1 and 2). Such direct repeats are a characteristic feature of all retroposons and probably result from a staggered chromosome break at the site of insertion (Weiner et al., 1986).

3.2 Comparison of *ψH2B* with other pseudogenes

Most of the known retropseudogenes of the protein-coding genes contain remnants of a poly(A) tail in their 3' part (Vanin, 1985; Weiner et al., 1986; Maestre et al. 1995). Even retroposons, derived from normally non-adenylated RNAs, such as 7SL RNA or snRNAs usually contain an A-rich sequence at their 3' end (Weiner et al., 1986). It was suggested that a poly(A)-tail is necessary for priming of the reverse transcription (Weiner et al., 1986). However, this was apparently not the case for the *D. melanogaster ψH2B* gene. There is no A-stretch present in its 3' part, which is in agreement with the fact that replication-dependent histone mRNAs in the animal kingdom are not polyadenylated (Birnstiel, 1985). The homology between *ψH2B* and the H2B gene stops one nucleotide upstream of the 3' terminal stem-loop sequence (Fig. 3). An attractive hypothesis is that the 3' terminal stem-loop sequence has served as a primer for reverse transcription event, which resulted in the generation of the *ψH2B* gene (see Fig. 3).

So far no retropseudogenes of replication-dependent histone genes were found in animal kingdom. Numerous replication-dependent histone pseudogenes, as found in different species, are usually present in clusters together with functional histone genes and their homology to the functional genes extends

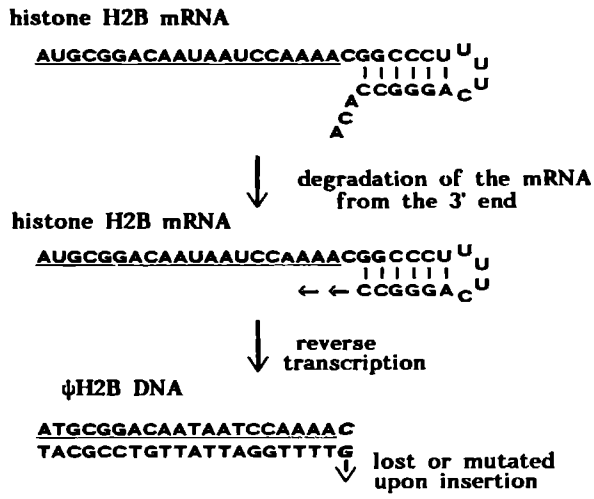


Fig. 3. Putative mechanism of generation of the *ψH2B* gene through a histone H2B mRNA intermediate. The part of the histone H2B 3'UTR, present in the pseudogene, is underlined.

beyond of the transcribed region (Turner et al., 1983; Marashi et al., 1984; Liu et al., 1987; Yang et al.,1987; Gruber et al., 1990; Kardalinou et al., 1993; Drabent et al., 1993; Dong et al., 1994; Brown et al., 1996). These pseudogenes apparently arose by mutational inactivation of the histone gene copies. The absense of retropseudogenes of replication-dependent histones was attributed to the lack of poly(A) tail in their mRNAs (Vanin, 1985). Our data on the *D. melanogaster ψH2B* gene show that the situation is not that simple and that replication-dependent histone genes can also be a source of retropseudogenes.

3.3. *ψH2B* was formed recently

The comparison of the *ψH2B* sequence with that of the normal H2B gene showed that in addition to the two point mutations, mentioned above, there are also other 4 non-synonymous and 6 synonymous substitutions in the coding region (see Fig. 1B). All the substitutions are transitions. The number of synonymous substitutions per site (K_S= 0.08) is considerably higher than the number of non-synonymous substitutions per site (K_n=0.02). This could indicate that *ψH2B* might have been under selection for some time prior to its inactivation (Miyata and Yasunaga, 1981). Using the average rate of synonymous substitution for *D. melanogaster* (16×10^{-9}/site/year, Moriyama and Gojobory, 1992), it can be estimated that the age of *ψH2B* is about 2.5 MY. This is probably an overestimate, because pseudogenes evolve faster than synonymous sites in functional genes. For mammalian globin genes the rate of substistution per site in pseudogenes is 1.9 times higher than the synonymous substitution rate in functional

genes (Miyata and Hayashida, 1981). If similar relationship is true for *Drosophila*, the real age of *ψH2B* is more close to 1.5 MY. Because of such a recent origin of the *ψH2B* gene, it can be predicted that it is not present even in the sibling species of *D. melanogaster* (divergence time with *D. melanogaster* is 2.5 MY, see Throckmorton, 1975).

3.4. *ψH2B* is surrounded by different transposons

Analysis of the sequences upstream and downstream of *ψH2B* revealed the presence of several transposable elements (Fig. 1 and 2). A ~500 bp sequence upstream of the *ψH2B* is 97% similar to the external half of the terminal inverted repeat of the foldback (*FB*) transposon, *FB4* (Potter, 1982). Downstream of *ψH2B* there is a region of ~450 bp, which is 60-80% similar to a number of *Drosophila* sequences, including the 3' UTR of the *male-specific lethal-2*, flanking regions of *suppressor of forked*, *beta subunit of integrin* and U1 snRNA genes 82.1 and 82.3. A small part of this sequence was identified as a repetitive element, which was called *Lefka* (Madueño et al., 1995). It seems likely that the whole 450 bp region represents a more complete copy of this repetitive element (Bashaw and Baker, 1995).

The regions upstream of *FB* and downstream of *Lefka* display some similarity to *MDG1* and *412* retrotransposons. The similarities are only obvious when the encoded proteins are compared (see Fig. 2). The sequences 5' of *FB* and sequences 3' of *Lefka* are similar to the middle part and the C-terminus, respectively, of the *MDG1* ORF 450 and of the ORF1 from the *412* retrotransposon. Still more downstream a region of similarity to the protease part of the ORF2 of the *412* is present. Such sequence organization is similar to that found in the *412* retrotransposon (Yuki et al., 1986). Large number of substitutions, resulting in frameshifts and stop codons, suggests that these sequences represent an inactive, degenerating copy of an *MDG1* - like retrotransposon.

No similarities to *MDG1* were found in the sequences between *FB* and *ψH2B* and between *ψH2B* and *Lefka*. It seems likely, therefore, that the combination of *FB-ψH2B-Lefka* was inserted into the retrotransposon as a result of a single event. It is known that *FB* elements can mobilize large DNA regions (see Lovering et al., 1991, and references therein). It is reasonable to propose that *ψH2B* and *Lefka* were "passengers" of a *FB* element, which was inserted into the degenerating *MDG1*-like retroposon. Subsequent loss of the right inverted repeat of the *FB* element would result in the observed arrangement of transposons.

Tight association of retropseudogenes with different transposable elements is not uncommon (see, for example, Wells and Bains, 1991; Koller et al., 1991). This is compatible with the idea that certain genomic regions represent hot spots for transposition.

Acknowledgements. We are grateful to Petra Bindels, Wenli Gu, Harry Harhangi and Koos Miedema for help and valuable discussions. We thank Koos Miedema for critical comments on the manuscript.

Chapter IX

General discussion

Anna Akhmanova

The work presented in this thesis has answered several questions, concerning the structure, organization and expression of replacement histone genes in *Drosophila*. Some new features of the *Drosophila* replication-dependent histone gene expression have also been described. In this chapter I will shortly discuss these results in the context of possible future experiments.

1. Does *Drosophila* have a full complement of replacement histone genes?

The replication-dependent genes, encoding all five histones, were isolated from many different organisms, while no full set of replacement genes was described so far in any organism. Genes, coding for replacement histones H1, H2A and H3 were found in many species (see Chapter I). In *Drosophila*, replacement genes were described for histone H2A (*H2AvD*, van Daal et al., 1988), H3 (*H3.3A* and *H3.3B*, this thesis, Chapter II) and H4 (*H4r*, this thesis, Chapter VII). The question arises if the replacement genes for histones H2B and H1 are also present in the fly genome.

A replacement histone H2B variant was never described in any species. However, our finding, that the replacement gene for histone H4 exists and encodes a protein, identical to the cell-cycle regulated H4, shows that a replacement gene does not necessarily encode a protein variant, detectable as a separate band on protein gels. It is quite possible that a replacement histone H2B gene indeed exists, but encodes a protein, very similar or even identical to the cell-cycle regulated histone H2B. Northern blots with *D. hydei* RNA support this idea, because polyadenylated RNA species similar, but not identical in sequence to the cell-cycle regulated histone H2B probe, can be detected (see Kremer, 1991; this thesis, Fig. 5 in Chapter IV). This is exactly what one would expect for transcripts, generated from a replacement histone H2B gene. However, our attempt to isolate such a gene from a *D. melanogaster* genomic DNA library by screening for an H2B-homologous sequence, unlinked to histone H4 genes, has failed. We have isolated a histone H2B pseudogene, but not a replacement histone H2B gene (see Chapter VIII). This failure does not prove the absense of the H2B repacement gene in the fly genome, since a single library usually does not cover 100% of the genome. It seems worthwhile to use the same approach for screening of a different genomic library. Alternatively, one could try to isolate a cDNA for the replacement H2B gene, by using RT-PCR with degenerated histone H2B primers on polyadenylated mRNA from flies. This strategy proved to be very efficient for isolation of the histone *H4r* cDNA (Chapter VII).

Replacement genes for linker histones were found both in vertebrates and invertebrates (Chapter I). Therefore, it seems quite likely that such gene(s) are also present in *Drosophila*. Since these genes undergo relatively rapid evolution, they are difficult to isolate by hybridization screening. The best approach is probably to use RT-PCR with degenerated primers, targeted against conserved regions in the globular domain of histone H1 (Schulze and Schulze, 1995).

It should be noted that the effort to isolate replacement histone genes might be made unnecessary by the advancement of the *Drosophila* genome project.

Systematic sequencing of the genomes of other higher eukaryotes (man, *Caenorhabditis elegans*, *Arabidopsis thaliana*) will also eventually provide an answer to the question about the existence of a full set of replacement genes in these organisms and the structure of these genes.

2. What is the function of replacement histone genes?

Our interest in replacement histones in *Drosophila* was based on the fact that this organism provides good opportunities for genetic analysis. The isolation and characterization of the structure and localization of the *H3.3A*, *H3.3B* and *H4r* genes provides the necessary basis for searching for flies with mutations in these genes.

Deletion in the *Drosophila* H2A replacement gene, *H2AvD*, is recessive lethal (van Daal and Elgin, 1992). It reasonable to expect, therefore, that mutations in the *H4r* gene and the H3.3-encoding genes will also have a phenotype.

Searching for a mutant in histone *H4r* gene is facilitated by the fact that there is a fly strain with a P-element insertion only 150 bp upstream from the *H4r* transcription start site. Unfortunately, this P-element is inserted into the adjacent gene, *punt*, which encodes an essential receptor protein, and the insertion is recessive lethal (Ruberte et al., 1995). However, this P-element insertion can be used for generating deletions over the histone *H4r* gene by mobilization of the P-element. These deletions might be helpful for establishing the histone *H4r* function when introduced into the flies, bearing in their genome a rescue construct for the *punt* gene. The latter should not be a problem, since the rescue of the mutation in *punt* gene with P-element constructs was already described (Ruberte et al., 1995; Letsou et al., 1995). Deletions in the histone *H4r* gene might also be useful for subsequent genetic screens for mutations in this gene.

Our preliminary data on mobilization of the P-element in the *punt* gene show that a high proportion (20%) of the males, which have lost the P-element, are sterile (Akhmanova, unpublished). The sterility is dominant, which is in agreement with earlier data about the existence in this locus of the chromosome 3 (88D) of a region, haploinsufficient for male fertility (see Green et al., 1990 and references therein). The sterility is not a result of the mutation in the *punt* gene itself, since the *punt* allele with the P-element insertion, which otherwhise behaves as a null allele, is not dominant male sterile (Ruberte et al., 1995).

The sterile males, which we obtained, produce no motile sperm (Akhmanova and Hennig, unpublished). It seems likely that these flies have a mutation in the gene *whirligig*, which was identified genetically in the same genomic region (Green et al., 1990). This locus is haploinsufficient for male fertility and the mutant males do not produce motile sperm (Green et al., 1990). The high frequency of such mutations upon mobilization of the P-element in the *punt* gene suggests that *whirligig* may be tightly linked to *punt*. It is unlikely, however that *whirligig* is histone *H4r*, since it was shown genetically that the product of

whirligig interacts with tubulin-encoding genes. It is hard to imagine that histone H4 would interact with tubulin, although such a possibility can not be completely excluded. It was shown, for example, that in sea urchin sperm flagella, axonemal microtubules are stabilized by a protein, identical to histone H1 (Multigner et al., 1992).

Genetic analysis of the histone H3.3 function is complicated by the presence of two non-allelic genes, *H3.3A* and *H3.3B*, which encode exactly the same protein. The expression patterns of these two genes partially overlap and their functions, therefore, might be redundant to a certain extent. To elucidate the function of the H3.3 protein it would be necessary to knock out both genes. Finding a mutant in the *H3.3A* gene should not be too difficult, since the region, where it is localized (25C), was extensively characterized genetically by Szidonya and Reuter (1988).

Searching for the H3.3B mutant is more complicated, since the region, where it is localized (9D) is poorly characterized genetically. We have checked all the available P-element insertion strains, containing P-elements in the region 9A-9F (Akhmanova, unpublished). In none of them a P-element was found to be inserted into or on the close vicinity of the histone H3.3B gene.

In spite of the difficulties, mentioned above, searching for the mutants in the described histone replacement genes is the most promising line of research, since it will hopefully provide the answers about the function of these genes.

3. How are replacement histones distributed in the chromosomes?

In order to analyse the chromosome distribution of the histone H3.3 variant we have raised an H3.3-specific antiserum (Chapter VI). We have demonstrated that at least in male germ line cells histone H3.3 is distributed in chromatin in a non-random fashion, which supports the idea that it has certain specific functions in these cells. It would be interesting to apply our serum in other systems, such as mammalian cells. This is possible, because the histone H3.3 variant is identical in all animals. Anti-H3.3 serum will also be very useful for the analysis of the mutants in the H3.3 genes, when these mutants are obtained.

A considerable limitation in the application of our anti-H3.3 serum is the necessity to remove the DNA from the cytological preparations. As was extensively discussed in Chapter VI, the removal of DNA might lead to artifacts. An alternative approach is to tag the replacement histones with an epitope, such as the influenza hemagglutinine epitope, hexahistidine tag or even green fluorescence protein. Tagged histone proteins can be tolerated by the cell and are incorporated into chromatin, as was shown for the yeast histone H2B (Lorch and Kornberg, 1994) and the centromeric histone H3-like protein, CENP-A (Sullivan et al., 1994). This approach can be used not only for the histone H3.3, but also for the histone *H4r* gene. For the latter it is the only possible way to discriminate its product from the product of the cell-cycle regulated histone H4 gene.

4. What is the function of the alternative polyadenylation and the long 3' UTR of the histone *H3.3B* gene?

D. melanogaster histone *H3.3B* gene produces transcripts with long 3'UTRs, which are subject to tissue-specific alternative polyadenylation (Chapter II). We have identified the exact positions of the alternatively used polyadenylation sites (Chapter II). As an initial step to characterize these polyadenylation sites, we compared their efficiency in the in vitro polyadenylation system from HeLa cells (Chapter III). Our results form a basis for the study of the mechanism of alternative polyadenylation of the H3.3B mRNAs. However, this mechanism can be elucidated only in an *Drosophila* in vitro polyadenylation system. Until now, such a system has not been developed.

The function of the long 3'UTR in the H3.3B gene expression can be addressed by making reporter gene constructs, containing a full-length 3'UTR or some parts of it. The sequence regions, conserved between *D. melanogaster* and *D. hydei* (see Chapter II), are good candidates for important regulatory elements. It would be interesting to know if their deletion has some impact on the reporter gene expression.

5. What are the specific features of histone gene expression in the male germ line cells?

While studying the various histone genes, described in this thesis, we have paid particular attention to their expression in *Drosophila* testes. As was mentioned in previous chapters, during male germ cell development in many organisms, chromatin undergoes severe rearrangements, which are accompanied by a complex pattern of expression of histone genes. Our data have shown that this conclusion holds true for *Drosophila*.

Some, but not all replication-dependent histone genes produce in addition to the expected short non-polyadenylated mRNAs longer polyadenylated transcripts, which are generated because of utilization of polyadenylation signals downstream of the stem-loop sequence. Such polyadenylated transcripts accumulate preferentially in the testis, where they might be used for synthesis of histone proteins after the cessation of DNA replication (Chapter IV).

Certain replacement histone genes, such as *H3.3A* (Chapters II and V) and *H4r* (see Fig. 1 in Chapter V), also display strong testis expression. For the histone *H3.3A* gene we have shown that it is strongly transcribed both in spermatogonia and primary spermatocytes (Chapter V). Concomitantly with the presence of the *H3.3A* mRNA in the cytoplasm of primary spermatocytes, we observed antisense transcripts of the cell-cycle regulated histone H3 gene in spermatocyte nuclei (Chapter V). With the antibody, specific for the histone H3.3 variant, we could detect the presence of the histone H3.3 protein in the spermatocyte nuclei and all postmeiotic stages until post-elongation spermatids (Chapter VI). One possible interpretation of these observations is that the replication-dependent histone H3 is partially substituted by the histone H3.3 in primary spermatocytes. Such sub-

stitution might require silencing of the replication-dependent histone H3 genes, which could be achieved by an antisense RNA-mediated mechanism.

Analysis of the regulatory functions of the histone H3 antisese transcripts will not be easy, since *Drosophila* testis is a very difficult tissue for biochemical studies due to its small size and problems in isolating of male germ cells of a particular developmental stage. The best way is probably to use constructs with a reporter gene, inserted instead of the histone H3 ORF in the replication-dependent histone gene repeat. The role of the antisense transcript in testis expression of the reporter gene can be then analysed by selectively mutating the sequences, controlling only the transcription of the antisense RNA. However, the latter might be difficult since the promoter of the antisense transcript coincides with the 3' processing site of the sense transcript (see Chapter V).

Immunocytology with antisera against core histones on testis preparations with hydrolyzed DNA has shown that during late stages of spermatid development histones are removed from chromatin (Chapter VI). This is in agreement with earlier cytochemical data, demonstrating the appearance of arginine-rich proteins in the nuclei of post-elongation spermatids (Hauschteck-Jungen and Hartl, 1982). The substitution of histones by the arginine-rich proteins does not occur in males, homozygous for segregation distorter (*SD*) (Kettaneh and Hartl, 1976). It would be interesting to apply our procedure for detecting core histones in the spermatid chromatin to the male germ cells from *SD* males in order to see if some abnormalities in the histone displacement can be detected.

References

Akhmanova, A S , Bindels, P C , Xu, J , Miedema, K , Kremer H , and Hennig W (1995) Structure and expression of histone H3 3 genes in Drosophila melanogaster and Drosophila hydei Genome 38, 586-600

Akhmanova, A , Miedema, K , and Hennig, W (1996) Identification and characterization of the Drosophila histone H4 replacement gene FEBS Lett 388, 219-222

Albig, W , Bramlage, B , Gruber, K , Klobeck, H G , Kunz, J , and Doenecke D (1995) The human replacement histone H3 3B gene (H3F3B) Genomics 30 264-272

Albig, W , Drabent, B , Kunz, J , Kalff-Suske, M , Grzeschik, K H , and Doenecke, D (1993) All known human H1 histone genes except the $H1^0$ gene are clustered on chromosome 6 Genomics 16, 649-654

Allis, C D , Richman, R , Gorovsky, M A , Ziegler, Y S , Touchstone, B Bradley W A , and Cook, R G (1986) hv1 is an evolutionarily conserved H2A variant that is preferentially associated with active genes J Biol Chem 261 1941 1948

Allis, C D , Ziegler, Y S , Gorovsky, M A , and Olmsted, J B (1982) A conserved histone variant enriched in nucleoli of mammalian cells Cell 31 131-136

Alt, F W , Bothwell, A L , Knapp, M , Siden, E Mather E , Koshland, M and Baltimore, D (1980) Synthesis of secreted and membrane-bound immunoglobulin mu heavy chains is directed by mRNAs that differ at their 3' ends Cell 20, 293-301

Anderson, K V and Lengyel, J A (1984) Histone gene expression in Drosophila development multiple levels of gene regulation In Histone genes Structure, organization and regulation, G S Stein, J L Stein, and W F Marzluff, eds (New York John Wiley and Sons), pp 135-161

Annunziato, A T , Eason, M B , and Perry, C A (1995) Relationship between methylation and acetylation of arginine-rich histones in cycling and arrested HeLa cells Biochemistry 34, 2916-2924

Arents, G , Burlingame, R W , Wang, B C , Love, W E , and Moudrianakis, E N (1991) The nucleosomal core histone octamer at 3 1 A resolution a tripartite protein assembly and a left-handed superhelix Proc Natl Acad Sci USA 88, 10148-10152

Arents, G and Moudrianakis, E N (1993) Topography of the histone octamer surface repeating structural motifs utilized in the docking of nucleosomal DNA Proc Natl Acad Sci USA 90, 10489-10493

Ashburner, M (1989) Drosophila a laboratory manual Cold Spring Harbor Laboratory Press

Ballantine, J E and Woodland, H R (1985) Polyadenylation of histone mRNA in Xenopus oocytes and embryos FEBS Lett 180, 224-228

Bannon, G A , Calzone, F J , Bowen, J K , Allis, C D , and Gorovsky, M A (1983) Multiple, independently regulated, polyadenylated messages for histone H3 and H4 in Tetrahymena Nucleic Acids Res 11, 3903-3917

Bannon, G A , and Gorovsky, M A (1984) Tetrahymena histones proteins messages and genes In Histone genes Structure, organization and regulation, G S Stein J L Stein, and W F Marzluff, eds (New York John Wiley and Sons), pp 163-179

Barratt, M J , Hazzalin, C A , Cano, E , and Mahadevan, L C (1994) Mitogen-stimulated phosphorylation of histone H3 is targeted to a small hyperacetylation-sensitive fraction Proc Natl Acad Sci USA 91, 4781-4785

Bashaw, G J and Baker, B S (1995) The msl-2 dosage compensation gene of Drosophila encodes a putative DNA-binding protein whose expression is sex specifically regulated by Sex-lethal Development 121, 3245-3258

Baxevanis A D and Landsman, D (1996) Histone Sequence Database a compilation of highly-conserved nucleoprotein sequences Nucleic Acids Res 24 245-247

Bedford, M , Arman, E , Orr-Urtreger, A and Lonai, P (1995) Analysis of the Hoxd-3 gene structure and localization of its sense and natural antisense transcripts DNA Cell Biol 14, 295-304

Bhatnagar Y M , McCullar, M K , and Chronister, R B (1984) Immunocytochemical localization of a histone H2A variant in the mammalian nucleolar chromatin Cell Biol Int Rep 8, 971-979

Bhatnagar, Y M , Faulker, R D , and McCullar, M K (1985) Biochemical and immunological characterization of an H2A variant from the mouse testis Biochim Biophys Acta 827, 14-22

Biard-Roche, J , Gorka, C , and Lawrence, J J (1982) The structural role of histone H1 properties of reconstituted chromatin with various H1 subfractions (H1-1, H1-2, and H1^{0}) EMBO J 1, 1487-1492

Birnstiel, M L , Busslinger, M , and Strub, K (1985) Transcription termination and 3' processing the end is in site Cell 41, 349-359

Boelens, W C , Jansen, E J , van Venrooij, W J , Stripecke, R , Mattaj, I W , and Gunderson S I (1993) The human U1 snRNP-specific U1A protein inhibits polyadenylation of its own pre-mRNA Cell 72, 881-892

Bonner, W M , Mannironi, C , Orr, A , Pilch, D R , and Hatch, C L (1993) Histone H2A X gene transcription is regulated differently than transcription of other replication-linked histone genes Mol Cell Biol 13, 984-992

Boulikas, T (1985) Electrophoretic separation of histones and high-mobility-group proteins on acid-urea-Triton gels Anal Biochem 149, 379-386

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

Bowman, T L and Hurt, M M (1995) The coding sequences of mouse H2A and H3 histone genes contains a conserved seven nucleotide element that interacts with nuclear factors and is necessary for normal expression Nucleic Acids Res 23 3083-3092

Bradbury, E M (1992) Reversible histone modifications and the chromosome cell cycle Bioessays 14, 9-16

Brawerman, G (1981) The role of the poly(A) sequence in mammalian RNA CRC Crit Rev Biochem 10, 1-38

Breathnach, R and Chambon, P (1981) Organization and expression of eucaryotic split genes coding for proteins Ann Rev Biochem 50, 349-383

Briand, J P , Barin, C , Van Regenmortel, M H , and Muller, S (1992) Application and limitations of the multiple antigen peptide (MAP) system in the production and evaluation of anti-peptide and anti-protein antibodies J Immunol Methods 156 255-265

Brown, D T , Wellman, S E , and Sittman, D B (1985) Changes in the levels of three different classes of histone mRNA during murine erythroleukemia cell differentiation Mol Cell Biol 5, 2879-2886

Brown, D T , Alexander, B T , and Sittman, D B (1996) Differential effect of H1 variant overexpression on cell cycle progression and gene expression Nucleic Acids Res 24, 486-493

Brown, V D , Wang, Z F , Williams, A S , and Marzluff, W F (1996) Structure of a cluster of mouse histone genes Biochim Biophys Acta 1306, 17-22

Brush, D , Dodgson, J B , Choi, O R , Stevens, P W , and Engel, J D (1985) Replacement variant histone genes contain intervening sequences Mol Cell Biol 5 1307-1317

Busslinger, M and Barberis, A (1985) Synthesis of sperm and late histone cDNAs of the sea urchin with a primer complementary to the conserved 3 terminal palindrome evidence for tissue-specific and more general histone gene variants Proc Natl Acad Sci USA 82, 5676-5680

Carr A M , Dorrington, S M , Hindley, J , Phear, G A , Aves, S J , and Nurse, P (1994) Analysis of a histone H2A variant from fission yeast Evidence for a role in chromosome stability Mol Gen Genet 245, 628-635

Chaboute, M E , Chaubet, N Gigot, C , and Philipps, G (1993) Histones and histone genes in higher plants structure and genomic organization Biochimie 75, 523-531

Challoner, P.B., Moss, S.B., and Groudine, M. (1989). Expression of replication-dependent histone genes in avian spermatids involves an alternate pathway of mRNA 3'-end formation. Mol. Cell. Biol. 9, 902-913.

Chalmers, M. and Wells, D. (1990). Extreme sequence conservation characterizes the rabbit H3.3A histone cDNA. Nucleic Acids Res. 18, 3075.

Chaubet, N., Clement, B., and Gigot, C. (1992). Genes encoding a histone H3.3-like variant in Arabidopsis contain intervening sequences. J. Mol. Biol. 225, 569-574.

Chen, C., Malone, T., Beckendorf, S. K. and Davis, R. L. (1987). At least two genes reside within a large intron of the dunce gene of Drosophila. Nature 329, 721-724.

Cheng, G.H., Nandi, A., Clerk, S., and Skoultchi, A.I. (1989). Different 3'-end processing produces two independently regulated mRNAs from a single H1 histone gene. Proc. Natl. Acad. Sci. USA 86, 7002-7006.

Childs, G., Maxson, R., Cohn, R.H., and Kedes, L. (1981). Orphons: dispersed genetic elements derived from tandem repetitive genes of eucaryotes. Cell 23, 651-663.

Chipev, C.C. and Wolffe, A.P. (1992). Chromosomal organization of Xenopus laevis oocyte and somatic 5S rRNA genes in vivo. Mol. Cell. Biol. 12, 45-55.

Chirgwin, J.M., Przybyla, A.E., MacDonald, R.J. and Rutter, W.J. (1979). Isolation of biologically active ribonucleic acid from sources enriched in ribonuclease. Biochemistry 18, 5294-5299.

Cho, H. and Wolffe, A.P. (1994). Xenopus laevis B4, an intron-containing oocyte-specific linker histone-encoding gene. Gene 143, 233-238.

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.

Cole, R.D. (1987). Microheterogeneity in H1 histones and its consequences. Int. J. Pept. Protein Res. 30, 433-449.

Coles, L.S. and Wells, J.R. (1985). An H1 histone gene-specific 5' element and evolution of H1 and H5 genes. Nucleic Acids Res. 13, 585-594.

Contursi, C., Minchiotti, G. and Dinocera, P. P. (1993). Functional dissection of two promoters that control sense and antisense transcription of Drosophila melanogaster F-elements. J. Mol. Biol. 234, 988-997.

D'Incalci, M., Allavena, P., Wu, R.S., and Bonner, W.M. (1986). H1 variant synthesis in proliferating and quiescent human cells. Eur. J. Biochem. 154, 273-279.

Das, C.C., Kaufmann, B.P. and Gay, H. (1964). Histone-protein transition in Drosophila melanogaster. Exp. Cell Res. 35, 507-514.

Dasso, M., Dimitrov, S., and Wolffe, A.P. (1994). Nuclear assembly is independent of linker histones. Proc. Natl. Acad. Sci. USA 91, 12477-12481.

Davie, J.R. and Delcuve, G.P. (1991). Characterization and chromatin distribution of the H1 histones and high-mobility-group non-histone chromosomal proteins of trout liver and hepatocellular carcinoma. Biochem. J. 280, 491-497.

Delcuve, G.P. and Davie, J.R. (1989). Chromatin structure of erythroid-specific genes of immature and mature chicken erythrocytes. Biochem. J. 263, 179-186.

Dell'Orco, R.T. and Worthington, M.I. (1988). The effects of in vitro age and culture state on histone variant synthesis in human diploid fibroblasts. J. Cell. Physiol. 136, 168-174.

Dijkman, H. B. P. M., Mentzel, S., de Jong, A. S. and Assmann, K. J. M. (1995) RNA in situ hybridization using digoxigenin-labeled cRNA probes. Biochemica N2, 21-25.

Dimitrov, S., Almouzni, G., Dasso, M., and Wolffe, A.P. (1993). Chromatin transitions during early Xenopus embryogenesis: changes in histone H4 acetylation and in linker histone type. Dev. Biol. 160, 214-227.

Dimitrov, S., Dasso, M.C., and Wolffe, A.P. (1994). Remodeling sperm chromatin in Xenopus laevis egg extracts: the role of core histone phosphorylation and linker histone B4 in chromatin assembly. J. Cell Biol. 126, 591-601.

Doenecke D and Tonjes, R (1986) Differential distribution of lysine and arginine residues in the closely related histones H1 and H5 Analysis of a human H1 gene J Mol Biol 187, 461-464

Doenecke, D , Albig, W , Bouterfa, H , and Drabent, B (1994) Organization and expression of H1 histone and H1 replacement histone genes J Cell Biochem 54, 423-431

Doenecke, D and Alonso, A (1996) Organization and expression of the developmentally regulated $H1^0$ histone gene in vertebrates Int J Dev Biol 40, 395-401

Dolnick, B J (1993) Cloning and characterization of a naturally occurring antisense RNA to human thymidylate synthase mRNA Nucleic Acids Res 21, 1747-1752

Donahue, P R , Palmer, D K , Condie, J M , Sabatini, L M , and Blumenfeld, M (1986) Drosophila histone H2A 2 is associated with the interbands of polytene chromosomes Proc Natl Acad Sci USA 83, 4744-4748

Dong Y Sirotkin, A M , Yang, Y S , Brown, D T , Sittman, D B , and Skoultchi, A I (1994) Isolation and characterization of two replication-dependent mouse H1 histone genes Nucleic Acids Res 22, 1421-1428

Dorsett, D (1990) Potentiation of a polyadenylylation site by a downstream protein-DNA interaction Proc Natl Acad Sci USA 87, 4373-4377

Drabent B , Kardalinou E , and Doenecke, D (1991) Structure and expression of the human gene encoding testicular H1 histone (H1t) Gene 103, 263-26 8

Drabent B , Kunz, C , and Doenecke D (1993) A rat histone H2B pseudogene is closely associated with the histone H1d gene Biochim Biophys Acta 1172, 193-196

Dworkin-Rastl E , Kandolf, H , and Smith, R C (1994) The maternal histone H1 variant, H1M (B4 protein), is the predominant H1 histone in Xenopus pregastrula embryos Dev Biol 161, 425-439

Eckner, R Ellmeier, W and Birnstiel, M L (1991) Mature mRNA 3' end formation stimulates RNA export from the nucleus EMBO J 10, 3513-3522

Elledge, S J and Davis R W (1989) Position and density effects on repression by stationary and mobile DNA binding proteins Genes Dev 3, 185-197

Enriquez-Harris, P , Levitt, N , Briggs, D , and Proudfoot, N J (1991) A pause site for RNA polymerase II is associated with termination of transcription EMBO J 10, 1833-1842

Fabry, S Muller, K , Lindauer, A , Park, P B Cornelius, T , and Schmitt, R (1995) The organization structure and regulatory elements of Chlamydomonas histone genes reveal features linking plant and animal genes Curr Genet 28, 333-345

Franklin, S G and Zweidler, A (1977) Non-allelic variants of histones 2a, 2b and 3 in mammals Nature 266, 273-275

Fretzin S Allan B D van Daal A , and Elgin, S C (1991) A Drosophila melanogaster H3 3 cDNA encodes a histone variant identical with the vertebrate H3 3 Gene 107, 341-342

Frohman M A (1990) RACE rapid amplification of cDNA ends In PCR protocols A guide to methods and applications, Innis, M A , Gelfand, D H , Sninsky, J J and White T J eds (San Diego Academic Press), pp 28-38

Fucci, L , Aniello F , Branno M Biffali, E , and Geraci, G (1994) Isolation of a new H3 3 histone variant cDNA of P lividus sea urchin sequence and embryonic expression Biochim Biophys Acta 1219, 539-542

Gjerset R , Gorka, C , Hasthorpe, S , Lawrence, J J , and Eisen, H (1982) Developmental and hormonal regulation of protein $H1^0$ in rodents Proc Natl Acad Sci USA 79 2333 2337

Glaser, R L and Lis J T (1990) Multiple, compensatory regulatory elements specify spermatocyte-specific expression of the Drosophila melanogaster hsp26 gene Mol Cell Biol 10, 131-137

Glatzer K H (1984) Preservation of nuclear RNP antigens in male germ cell development of Drosophila hydei Mol Gen Genet 196, 236-243

Gorka, C , Fakan, S , and Lawrence, J J (1993) Light and electron microscope immunocytochemical analyses of histone $H1^0$ distribution in the nucleus of Friend erythroleukemia cells Exp Cell Res 205, 152-158

Gorka, C , Lawrence, J J , and Khochbin, S (1995) Variation of $H1^0$ content throughout the cell cycle in regenerating rat liver Exp Cell Res 217, 528-533

Graves, R A , Pandey, N B , Chodchoy, N and Marzluff, W F (1987) Translation is required for regulation of histone mRNA degradation Cell 48, 615-626

Grimes, S R , Wolfe, S A , Anderson, J V , Stein, G S , and Stein, J L (1990) Structural and functional analysis of the rat testis-specific histone H1t gene J Cell Biochem 44, 1-17

Grond, C J (1984) Spermatogenesis in Drosophila Ph D thesis University of Nijmegen, the Netherlands

Grove, G W and Zweidler, A (1984) Regulation of nucleosomal core histone variant levels in differentiating murine erythroleukemia cells Biochemistry 23 4436-4443

Green, L L , Wolf, N , McDonald, K L , and Fuller, M T (1990) Two types of genetic interaction implicate the whirligig gene of Drosophila melanogaster in microtubule organization in the flagellar axoneme Genetics 126, 961-973

Gruber, A , Streit, A , Reist, M , Benninger, P , Bohni, R , and Schumperli D (1990) Structure of a mouse histone-encoding gene cluster Gene 95, 303-304

Grunstein, M , Rykowski, M , Kolodrubetz, D , Choe, J , and Wallis, J (1984) A genetic analysis of histone protein subtypes in yeast In Histone genes Structure, organization and regulation, G S Stein, J L Stein, and W F Marzluff, eds (New York John Wiley and Sons), pp 35-63

Grunstein, M , Durrin, L K , Mann, R K , Fischer-Adams, G , and Johnson, L M (1992) Histones regulators of transcription in yeast In Transcriptional regulation, S McKnight and K Yamamoto, eds (New York, Cold Spring Harbor Cold Spring Harbor Laboratory Press), pp 1295-1315

Grunwald, D , Khochbin, S , and Lawrence, J J (1991) Cell cycle-related accumulation of $H1^0$ mRNA induction in murine erythroleukemia cells Exp Cell Res 194, 174-179

Harlow, E and Lane, D (1988) Antibodies, a laboratory manual Cold Spring Harbor Laboratory

Hatch, C L and Bonner, W M (1988) Sequence of cDNAs for mammalian H2A Z, an evolutionarily diverged but highly conserved basal histone H2A isoprotein species Nucleic Acids Res 16, 1113-1124

Hatch, C L and Bonner, W M (1990) The human histone H2A Z gene Sequence and regulation J Biol Chem 265, 15211-15218

Hauschteck-Jungen, E and Hartl, D L (1982) Defective histone transition during spermiogenesis in heterozygous segregation distorter males of Drosophila melanogaster Genetics 101, 57-69

Hazelrigg, T and Tu, C (1994) Sex-specific processing of the Drosophila exuperantia transcript is regulated in male germ cells by the tra-2 gene Proc Natl Acad Sci USA 91, 10752-10756

Hecht, N B (1989a) Molecular biology of structural chromosomal proteins of the mammalian testis In Molecular biology of chromosome function (ed Adolph, K W), pp 396-420 Springer Verlag, New York

Hecht, N B (1989b) Mammalian protamines and their expression In Histones and other basic nuclear proteins (ed L Hnilica, G S Stein and J L Stein), pp 347-373 CRC, Boca Raton, Florida

Hedley, M L and Maniatis, T (1991) Sex-specific splicing and polyadenylation of dsx pre-mRNA requires a sequence that binds specifically to tra-2 protein in vitro Cell 65, 579-586

Helfman, D M , Cheley, S , Kuismanen, E , Finn, L A , and Yamawaki Kataoka Y (1986) Nonmuscle and muscle tropomyosin isoforms are expressed from a single gene by alternative RNA splicing and polyadenylation Mol Cell Biol 6, 3582-3595

Hendzel, M J and Davie, J R (1990) Nucleosomal histones of transcriptionally active/competent chromatin preferentially exchange with newly synthesized histones in quiescent chicken erythrocytes Biochem J 271, 67-73

Hennig, W (1967) Untersuchungen zur Struktur und Funktion des Lampenbürsten-Y-Chromosoms in der Spermatogenese von Drosophila Chromosoma 22, 294-357

Hennig, W (1987) The Y chromosomal lampbrush loops of Drosophila In Structure and function of eukaryotic chromosomes W Hennig, ed Springer-Verlag, Berlin, pp 133-146

Hennig, W , Brand, R C , Hackstein, J , Hochstenbach, R , Kremer, H , Lankenau, D H , Lankenau, S , Miedema, K , and Potgens, A (1989) Y chromosomal fertility genes of Drosophila a new type of eukaryotic genes Genome 31, 561-571

Hennig, W and Kremer, H (1990) Spermatogenesis of Drosophila hydei Int Rev Cytol 123, 129-175

Henikoff, S , Keene, M A , Fechtel, K and Fristrom, J W (1986) Gene within a gene nested Drosophila genes encode unrelated proteins on opposite DNA strands Cell 44, 33-42

Hildebrandt, M and Nellen, W (1992) Differential antisense transcription from the Dictyostelium EB4 gene locus implications on antisense-mediated regulation of mRNA stability Cell 69, 197-204

Hoheisel, J D , Lennon, G G , Zehetner, G and Lehrach, H (1991) Use of high coverage reference libraries of Drosophila melanogaster for relational data analysis J Mol Biol 220, 903-914

Hraba-Renevey, S and Kress, M (1989) Expression of a mouse replacement histone H3 3 gene with a highly conserved 3 noncoding region during SV40- and polyoma-induced Go to S-phase transition Nucleic Acids Res 17, 2449-2461

Hsieh Li, H M , Witte, D P , Weinstein, M , Branford, W , Li, H , Small, K and Potter S S (1995) Hoxa 11 structure, extensive antisense transcription, and function in male and female fertility Development 121, 1373-1385

Ivanova, V S , Hatch, C L , and Bonner, W M (1994a) Characterization of the human histone H2A X gene Comparison of its promoter with other H2A gene promoters J Biol Chem 269, 24189-24194

Ivanova, V S , Zimonjic, D , Popescu, N , and Bonner, W M (1994b) Chromosomal localization of the human histone H2A X gene to 11q23 2-q23 3 by fluorescence in situ hybridization Hum Genet 94, 303-306

Jackson, R J (1993) Cytoplasmic regulation of mRNA function the importance of the 3 untranslated region Cell 74 9-14

Jackson R J , and Standart N (1990) Do the poly(A) tail and 3' untranslated region control mRNA translation? Cell 62, 15-24

Jasinskiene, N , Jasinskas, A , and Langmore, J P (1995) Embryonic regulation of histone ubiquitination in the sea urchin Dev Genet 16, 278-290

Kaludov, N K Pabon-Pena, L , and Hurt, M M (1996) Identification of a second conserved element within the coding sequence of a mouse H3 histone gene that interacts with nuclear factors and is necessary for normal expression Nucleic Acids Res 24, 523-531

Kanazin, V , Blake, T , and Shoemaker, R C (1996) Organization of the histone H3 genes in soybean, barley and wheat Mol Gen Genet 250, 137-147

Kapros, T , Bogre, L , Nemeth, K , Bako, L , Gyorgyey, J , Wu, S -C , and Dudits, D (1992) Differential expression of histone H3 gene variants during cell cycle and somatic embryogenesis in alfalfa Plant Physiol 98, 621-625

Kapros, T , Robertson, A J , and Waterborg, J H (1995) Histone H3 transcript stability in alfalfa Plant Mol Biol 28, 901-914

Kardalinou E , Eick, S , Albig, W , and Doenecke, D (1993) Association of a human H1 histone gene with an H2A pseudogene and genes encoding H2B 1 and H3 1 histones J Cell Biochem 52, 375-383

Kedes, L H (1979) Histone genes and histone messengers Annu Rev Biochem 48 837 870

Keller,W (1995) No end yet to messenger RNA 3' processing Cell 81 829-832

Kettaneh, N P and Hartl, D L (1976) Histone transition during spermiogenesis is absent in segregation distorter males of Drosophila melanogaster Science 193, 1020-1021

Khochbin, S and Lawrence, J J (1989) An antisense RNA involved in p53 mRNA maturation in murine erythroleukemia cells induced to differentiate EMBO J 8 4107-4114

Khochbin, S , Gorka, C , and Lawrence, J J (1991) Multiple control levels governing $H1^0$ mRNA and protein accumulation FEBS Lett 283, 65-67

Khochbin, S and Wolffe, A P (1994) Developmentally regulated expression of linker-histone variants in vertebrates Eur J Biochem 225, 501-510

Kimelman, D and Kirschner, M W (1989) An antisense mRNA directs the covalent modification of the transcript encoding fibroblast growth factor in Xenopus oocytes Cell 59, 687-696

Kindy, M S , McCormack, J E , Buckler, A J , Levine, R A and Sonenshein, G E (1987) Independent regulation of transcription of the two strands of the c-myc gene Mol Cell Biol 7, 2857-2862

Kirsh, A L , Groudine, M , and Challoner, P B (1989) Polyadenylation and U7 snRNP-mediated cleavage alternative modes of RNA 3' processing in two avian histone H1 genes Genes Dev 3, 2172-2179

Kleinschmidt, J A and Steinbeisser, H (1991) DNA-dependent phosphorylation of histone H2A X during nucleosome assembly in Xenopus laevis oocytes involvement of protein phosphorylation in nucleosome spacing EMBO J 10, 3043-3050

Koller, M , Baumer, A , and Strehler, E E (1991) Characterization of two novel human retropseudogenes related to the calmodulin-encoding gene, CaMII Gene 97, 245-251

Kozak, M (1991) An analysis of vertebrate mRNA sequences intimations of translational control J Cell Biol 115, 887-903

Kramers, C (1995) Anti-nucleosome auto-antibodies in systemic lupus erythematosus Ph D thesis University of Nijmegen, Nijmegen, The Netherlands

Kramers, C , Hylkema, M , Termaat, R M , Brinkman, K , Smeenk, R , and Berden, J H M (1993) Histones in lupus nephritis Exp Nephrol 1, 224-228

Kramers, C , Hylkema, M N , van Bruggen, M C , van de Lagemaat, R , Dijkman, H B , Assmann, K J , Smeenk, R J , and Berden, J H M (1994) Anti-nucleosome antibodies complexed to nucleosomal antigens show anti-DNA reactivity and bind to rat glomerular basement membrane in vivo J Clin Invest 94, 568-577

Kremer, H (1991) Spermatogenesis of Drosophila hydei chromatin reorganization and histone variants Ph D Thesis (The Netherlands Catholic University of Nijmegen)

Kremer, H , Hennig, W and Dijkhof, R (1986) Chromatin organization in the male germ line of Drosophila hydei Chromosoma 94, 147-161

Kremer, H and Hennig, W (1990) Isolation and characterization of a Drosophila hydei histone DNA repeat unit Nucleic Acids Res 18, 1573-1580

Krieg, P A , Robins, A J , D'Andrea, R , and Wells, J R (1983) The chicken H5 gene is unlinked to core and H1 histone genes Nucleic Acids Res 11, 619-627

Krimer, D B , Cheng, G , and Skoultchi, A I (1993) Induction of H3 3 replacement histone mRNAs during the precommitment period of murine erythroleukemia cell differentiation Nucleic Acids Res 21, 2873-2879

Kruger, W , Peterson, C L , Sil, A , Coburn, C , Arents, G , Moudrianakis, E N and Herskowitz, I (1995) Amino acid substitutions in the structured domains of histones H3 and H4 partially relieve the requirement of the yeast SWI/SNF complex for transcription Genes Dev 9, 2770-2779

Krystal, G W , Armstrong, B C and Battey, J F (1990) N-myc mRNA forms an RNA-RNA duplex with endogenous antisense transcripts Mol Cell Biol 10, 4180-4191

Kuhn, R., Kuhn, C., Borsch, D., Glatzer, K. H., Schafer, U. and Schafer, M. (1991). A cluster of four genes selectively expressed in the male germ line of Drosophila melanogaster. Mech. Dev. 35, 143-151.

Laabi, Y., Gras, M. P., Brouet, J. C., Berger, R., Larsen, C. J. and Tsapis, A. (1994). The BCMA gene, preferentially expressed during B lymphoid maturation, is bidirectionally transcribed. Nucleic Acids Res. 22, 1147-1154.

Lai, Z.C. and Childs, G. (1986). Isolation and characterization of the gene encoding the testis specific histone protein H2B-2 from the sea urchin Lytechinus pictus. Nucleic Acids Res. 14, 6845-6856.

Lai, Z., Lieber, T., and Childs, G. (1986). The nucleotide sequence of the gene encoding the sperm specific histone subtype H2B-1 from the sea urchin Strongylocentrotus purpuratus. Nucleic Acids Res. 14, 9218.

Lankenau, S., Corces, V. G. and Lankenau, D. H. (1994). The Drosophila micropia retrotransposon encodes a testis-specific antisense RNA complementary to reverse transcriptase. Mol. Cell. Biol. 14, 1764-1775.

Lennox, R.W. and Cohen, L.H. (1983). The histone H1 complements of dividing and nondividing cells of the mouse. J. Biol. Chem. 258, 262-268.

Lennox, R.W. and Cohen, L.H. (1989). Analysis of histone subtypes and their modified forms by polyacrylamide gel electrophoresis. Methods Enzymol. 170, 532-549.

Letsou, A., Arora, K., Wrana, J.L., Simin, K., Twombly, V., Jamal, J., Staehling-Hampton, K., Hoffmann, F.M., Gelbart, W.M., Massague, J., and O'Connor, M.B. (1995). Drosophila dpp signaling is mediated by the punt gene product: A dual ligand-binding type II receptor of the TGF beta receptor family. Cell 80, 899-908.

Levine, B.J., Chodchoy, N., Marzluff, W.F. and Skoultchi, A.I. (1987). Coupling of replication type histone mRNA levels to DNA synthesis requires the stem-loop sequence at the 3' end of the mRNA. Proc. Natl. Acad. Sci. USA. 84, 6189-6193.

Li, W., Nagaraja, S., Delcuve, G.P., Hendzel, M.J., and Davie, J.R. (1993). Effects of histone acetylation, ubiquitination and variants on nucleosome stability. Biochem. J. 296, 737-744.

Lifton, R.P., Goldberg, M.L., Karp, R.W., and Hogness, D.S. (1978). The organization of the histone genes in Drosophila melanogaster: functional and evolutionary implications. Cold Spring Harb. Symp. Quant. Biol. 42 Pt. 2, 1047-1051.

Lindsley, D. L. (1980). Spermatogenesis. In The genetics and biology of Drosophila (ed. M. Ashburner and T.R.F. Wright), pp. 225-294. Academic Press, London.

Lindsley, D.L. and Tokuyasu, K.T. (1980). Spermatogenesis. In; The genetics and biology of Drosophila, v. 2. M. Ashburner and T.R.F. Wright, eds. Academic Press, London pp. 225-294.

Liu, T.J., Liu, L., and Marzluff, W.F. (1987). Mouse histone H2A and H2B genes: four functional genes and a pseudogene undergoing gene conversion with a closely linked functional gene. Nucleic Acids Res. 15, 3023-3039.

Liu, X. and Gorovsky, M.A. (1993). Mapping the 5' and 3' ends of Tetrahymena thermophila mRNAs using RNA ligase mediated amplification of cDNA ends (RLM-RACE). Nucleic Acids Res. 21, 4954-4960.

Liu, X., Bowen, J., and Gorovsky, M.A. (1996). Either of the major H2A genes but not an evolutionarily conserved H2A.F/Z variant of Tetrahymena thermophila can function as the sole H2A gene in the yeast Saccharomyces cerevisiae. Mol. Cell. Biol. 16, 2878-2887.

Lorch, Y. and Kornberg, R.D. (1994). Isolation of the yeast histone octamer. Proc. Natl. Acad. Sci. USA 91, 11032-11034.

Lou, H., Gagel, R.F., and Berget, S.M. (1996). An intron enhancer recognized by splicing factors activates polyadenylation. Genes Dev. 10, 208-219.

Lovering, R., Harden, N., and Ashburner, M. (1991). The molecular structure of TE146 and its derivatives in Drosophila melanogaster. Genetics 128, 357-372.

Madueno, E., Papagiannakis, G., Rimmington, G., Saunders, R.D.C., Savakis, C., Sidenkiamos, I., Skavdis, G., Spanos, L., Trenear, J., Adam, P., Ashburner, M.,

Benos, P , Bolshakov, V N , Coulson, D , Glover, D M , Herrmann, S , Kafatos, F C , Louis, C , Majerus, T , and Modolell, J (1995) A physical map of the X chromosome of Drosophila melanogaster Cosmid contigs and sequence tagged sites Genetics 139, 1631-1647

Maestre, J , Tchenio, T , Dhellin, O , and Heidmann, T (1995) mRNA retroposition in human cells processed pseudogene formation EMBO J 14, 6333-6338

Mannironi, C , Bonner, W M , and Hatch, C L (1989) H2A X a histone isoprotein with a conserved C-terminal sequence, is encoded by a novel mRNA with both DNA replication type and polyA 3' processing signals Nucleic Acids Res 17, 9113-9126

Marashi, F , Prokopp, K , Stein, J , and Stein, G (1984) Evidence for a human histone gene cluster containing H2B and H2A pseudogenes Proc Natl Acad Sci

Marion, C , Roche, J , Roux, B , and Gorka, C (1985) Differences in the condensation of chromatin by individual subfractions of histone H1 implications for the role of $H1^0$ in the structural organization of chromatin Biochemistry 24, 6328-6335

Marzluff, W F (1992) Histone 3' ends essential and regulatory functions Gene Expr 2, 93-97

Marzluff, W F and Pandey, N B (1988) Multiple regulatory steps control histone mRNA concentrations Trends Biochem Sci 13, 49-52

Matsuo, Y and Yamazaki, T (1989) tRNA derived insertion element in histone gene repeating unit of Drosophila melanogaster Nucleic Acids Res 17, 225-238

Mendelson, E , Landsman, D , Druckmann, S , and Bustin, M (1986) Immunofractionation of chromatin regions associated with histone $H1^0$ Eur J Biochem 160, 253-260

Michiels, F , Wolk, A and Renkawitz-Pohl, R (1991) Further sequence requirements for male germ cell-specific expression under the control of the 14 bp promoter element (β2UE1) of the Drosophila β2 tubulin gene Nucleic Acids Res 19 4515-4521

Miedema, K , Hanske, M , Akhmanova, A , Bindels, P and Hennig, W (1995) Minor-myosin, a novel myosin isoform synthesized preferentially in Drosophila testis is encoded by the muscle myosin heavy chain gene Mech Dev 51, 67-81

Miyajima, N , Horiuchi, R , Shibuya, Y , Fukushige, S , Matsubara, K , Toyoshima, K and Yamamoto, T (1989) Two erbA homologs encoding proteins with different T3 binding capacities are transcribed from opposite DNA strands of the same genetic locus Cell 57, 31 39

Miyata, T and Hayashida, H (1981) Extraordinary high evolutionary rate of pseudogenes Evidence for the presence of selective pressure against changes between synonymous codons Proc Natl Acad Sci USA 78, 5739-5743

Miyata T and Yasunaga, T (1981) Rapidly evolving mouse α-globin-related pseudo gene and its evolutionary history Proc Natl Acad Sci USA 78, 450-453

Mohr E , Trieschmann, L , and Grossbach, U (1989) Histone H1 in two subspecies of Chironomus thummi with different genome sizes homologous chromosome sites differ largely in their content of a specific H1 variant Proc Natl Acad Sci USA 86, 9308-9312

Monestier, M and Kotzin, B L (1992) Antibodies to histones in systemic lupus erythematosus and drug-induced lupus syndromes Rheum Dis Clin North Am 18, 415-436

Moore, C L and Sharp, P A (1984) Site-specific polyadenylation in a cell-free reaction Cell 36, 581-591

Moore, C L and Sharp, P A (1985) Accurate cleavage and polyadenylation of exogenous RNA substrate Cell 41 845-855

Moriyama, E N and Gojobori T (1992) Rates of synonymous substitution and base composition of nuclear genes in Drosophila Genetics 130, 855-864

Moss, S B , Challoner P B and Groudine, M (1989) Expression of a novel histone 2B during mouse spermiogenesis Dev Biol 133, 83-92

Moss S B , Ferry, R A , and Groudine, M (1994) An alternative pathway of histone mRNA 3 end formation in mouse round spermatids Nucleic Acids Res 22, 3160-3166

Mowry, K L and Steitz, J A (1988) snRNP mediators of 3 end processing functional fossils? Trends Biochem Sci 13, 447-451

Mowry, K L , Oh, R and Steitz, J A (1989) Each of the conserved sequence elements flanking the cleavage site of mammalian histone pre-mRNAs has a distinct role in the 3'-end processing reaction Mol Cell Biol 9, 3105-3108

Muller, S , Chaix, M L , Briand, J P , and Van Regenmortel M H (1991) Immunogenicity of free histones and of histones complexed with RNA Mol Immunol 28, 763-772

Muller, S , Plaue, S , Couppez, M , and Van Regenmortel, M H (1986) Comparison of different methods for localizing antigenic regions in histone H2A Mol Immunol 23, 593-601

Müller, K and Schmitt, R (1988) Histone genes of Volvox carteri DNA sequence and organization of two H3-H4 gene loci Nucleic Acids Res 16, 4121-4136

Müller, K , Lindauer, A , Bruderlein, M and Schmitt, R (1990) Organization and transcription of Volvox histone-encoding genes similarities between algal and animal genes Gene 93, 167-175

Multigner, L , Gagnon, J , Van Dorsselaer, A , and Job, D (1992) Stabilization of sea urchin flagellar microtubules by histone H1 Nature 360, 33-39

Munroe, D and Jacobson, A (1990) Tales of poly(A) a review Gene 91, 151-158

Munroe, S H and Lazar, M A (1991) Inhibition of c-erbA mRNA splicing by a naturally occurring antisense RNA J Biol Chem 266, 22083-22086

Murashov, A and Wolgemuth, D J (1996) Distinct transcripts are recognized by sense and antisense riboprobes for a member of the murine HSP70 gene family, HSP70 2, in various reproductive tissues Mol Reprod Dev 43, 17-24

Nagaraja, S , Delcuve, G P , and Davie, J R (1995) Differential compaction of transcriptionally competent and repressed chromatin reconstituted with histone H1 subtypes Biochim Biophys Acta 1260, 207-214

Nagata, T , Kato, T , Morita, T , Nozaki, M , Kubota, H , Yagi, H , and Matsushiro, A (1991) Polyadenylated and 3' processed mRNAs are transcribed from the mouse histone H2A X gene Nucleic Acids Res 19, 2441-2447

Nei, M and Gojobori, T (1986) Simple methods for estimating the numbers of synonymous and nonsynonymous nucleotide substitutions Mol Biol Evol 3, 418-426

Nesic, D and Maquat, L E (1994) Upstream introns influence the efficiency of final intron removal and RNA 3'-end formation Gene Dev 8, 363-375

Nevins, J R and Wilson, M C (1981) Regulation of adenovirus-2 gene expression at the level of transcriptional termination and RNA processing Nature 290, 113-118

Nickel, B E , Roth, S Y , Cook, R G , Allis, C D , and Davie, J R (1987) Changes in the histone H2A variant H2A Z and polyubiquitinated histone species in developing trout testis Biochemistry 26, 4417-4421

Nomoto, M , Imai, N , Saiga, H , Matsui, T , and Mita, T (1987) Characterization of two types of histone H2B genes from macronuclei of Tetrahymena thermophila Nucleic Acids Res 15, 5681-5697

Norris, D and Osley, M A (1987) The two gene pairs encoding H2A and H2B play different roles in the Saccharomyces cerevisiae life cycle Mol Cell Biol 7, 3473-3481

Ohsumi, K , Katagiri, C , and Kishimoto, T (1993) Chromosome condensation in Xenopus mitotic extracts without histone H1 Science 262, 2033-2035

Osley, M A (1991) The regulation of histone synthesis in the cell cycle Annu Rev Biochem 60, 827-861

Pandey, N B and Marzluff, W F (1987) The stem-loop structure at the 3' end of histone mRNA is necessary and sufficient for regulation of histone mRNA stability Mol Cell Biol 7, 4557-4559

Pandey, N B , Chodchoy, N , Liu, T J , and Marzluff, W F (1990) Introns in histone genes alter the distribution of 3' ends Nucleic Acids Res 18, 3161-3170

Panyim, S. and Chalkley, R. (1969). A new histone found only in mammalian tissues with little cell division. Biochem. Biophys. Res. Commun. 37, 1042-1049.

Paranjape, S.M., Kamakaka, R.T., and Kadonaga, J.T. (1994). Role of chromatin structure in the regulation of transcription by RNA polymerase II. Annu. Rev. Biochem. 63, 265-297.

Pehrson, J.R. and Fried, V.A. (1992). MacroH2A, a core histone containing a large nonhistone region. Science 257, 1398-1400.

Pina, B. and Suau, P. (1987). Changes in histones H2A and H3 variant composition in differentiating and mature rat brain cortical neurons. Dev. Biol. 123, 51-58.

Pisano, C., Bonaccorsi, S. and Gatti, M. (1993). The kl-3 loop of the Y chromosome of Drosophila melanogaster binds a tektin-like protein. Genetics 133, 569-579.

Poccia, D. (1986). Remodeling of nucleoproteins during gametogenesis, fertilization, and early development. Int. Rev. Cytol. 105, 1-65.

Popescu, N., Zimonjic, D., Hatch, C., and Bonner, W. (1994). Chromosomal mapping of the human histone gene H2A.Z to 4q24 by fluorescence in situ hybridization. Genomics 20, 333-335.

Potter, S.S. (1982). DNA sequence of a foldback transposable element in Drosophila. Nature 297, 201-204.

Prescott, J. and Falck-Pedersen, E. (1994). Sequence elements upstream of the 3' cleavage site confer substrate strength to the adenovirus L1 and L3 polyadenylation sites. Mol. Cell. Biol. 14, 4682-4693.

Ramsey-Ewing, A., Van Wijnen, A.J., Stein, G.S., and Stein, J.L. (1994). Delineation of a human histone H4 cell cycle element in vivo: the master switch for H4 gene transcription. Proc. Natl. Acad. Sci. USA 91, 4475-4479.

Ridsdale, J.A. and Davie, J.R. (1987). Chicken erythrocyte polynucleosomes which are soluble at physiological ionic strength and contain linker histones are highly enriched in beta-globin gene sequences. Nucleic Acids Res. 15, 1081-1096.

Robertson, A.J., Kapros, T., Dudits, D., and Waterborg, J.H. (1996). Identification of three highly expressed replacement histone H3 genes of alfalfa. DNA Sequence 6, 137-146.

Roche, J., Girardet, J.L., Gorka, C., and Lawrence, J.J. (1985a). The involvement of histone $H1^0$ in chromatin structure. Nucleic Acids Res. 13, 2843-2853.

Roche, J., Gorka, C., Goeltz, P., and Lawrence, J.J. (1985b). Association of histone $H1^0$ with a gene repressed during liver development. Nature 314, 197-198.

Romano, G. (1992). Histone variants during sea urchin development. Cell Biol. Int. Rep. 16, 197-206.

Ross, J., Peltz, S.W., Kobs, G. and Brewer, G. (1986) Histone mRNA degradation in vivo: the first detectable step occurs at or near the 3' terminus, Mol. Cell. Biol. 6, 4362-4371.

Ruberte, E., Marty, T., Nellen, D., Affolter, M., and Basler, K. (1995). An absolute requirement for both the type II and type I receptors, punt and thick veins, for dpp signaling in vivo. Cell 80, 889-897.

Ruderman, J.V., Woodland, H.R. and Sturgess, E.A. (1979) Modulations of histone messenger RNA during the early development of Xenopus laevis, Dev. Biol. 71, 71-82.

Rungger-Brändle, E., Jamrich, M., and Bautz, E.K. (1981). Localization of RNA polymerase B and histones in the nucleus of primary spermatocytes of Drosophila hydei, studied by immunofluorescence microscopy. Chromosoma 82, 399-407.

Sachs, A.B. (1993). Messenger RNA degradation in eukaryotes. Cell 74, 413-421.

Schulz, R. A., Xie, X. L. and Miksch, J. L. (1990). cis-acting sequences required for the germ line expression of the Drosophila gonadal gene. Dev. Biol. 140, 455-458.

Schulze, E. and Schulze, B. (1995). The vertebrate linker histones $H1^0$, H5, and H1M are descendants of invertebrate "orphon" histone H1 genes. J. Mol. Evol. 41, 833-840.

Schümperli, D (1986) Cell-cycle regulation of histone gene expression Cell 45, 471-472

Schumperli, D (1988) Multilevel regulation of replication-dependent histone genes Trends Genet 4, 187-191

Sheets, M D , Ogg, S C , and Wickens, M P (1990) Point mutations in AAUAAA and the poly (A) addition site effects on the accuracy and efficiency of cleavage and polyadenylation in vitro Nucleic Acids Res 18, S799-S805

Shen, X , Yu, L , Weir, J W , and Gorovsky, M A (1995) Linker histones are not essential and affect chromatin condensation in vivo Cell 82, 47-56

Simpson, R T (1981) Modulation of nucleosome structure by histone subtypes in sea urchin embryos Proc Natl Acad Sci USA 78, 6803-6807

Simpson, R T and Bergman, L W (1980) Structure of sea urchin sperm chromatin core particle J Biol Chem 255, 10702-10709

Sirotkin, A M , Edelmann, W , Cheng, G , Klein Szanto, A , Kucherlapati, R , and Skoultchi, A I (1995) Mice develop normally without the $H1^0$ linker histone Proc Natl Acad Sci USA 92, 6434-6438

Sittman, D B , Graves, R A , and Marzluff, W F (1983) Histone mRNA concentrations are regulated at the level of transcription and mRNA degradation Proc Natl Acad Sci USA 80, 1849-1853

Smeenk, R J T ,Brinkman, K , Van den Brink, H ,G , and Westgeest, A ,A ,A (1988) Reaction patterns of monoclonal antibodies to DNA J Immunol 140, 3789 3792

Smith, R C , Dworkin-Rastl, E , and Dworkin, M B (1988) Expression of a histone H1-like protein is restricted to early Xenopus development Genes Dev 2, 1284-1295

Spencer, C A , Gietz, R D and Hodgetts, R B (1986) Overlapping transription units in the dopa carboxilase region of Drosophila Nature 322, 279-280

Stargell, L A , Bowen, J , Dadd, C A , Dedon, P C , Davis, M , Cook, R G , Allis, C D , and Gorovsky, M A (1993) Temporal and spatial association of histone H2A variant hv1 with transcriptionally competent chromatin during nuclear development in Tetrahymena thermophila Genes Dev 7, 2641-2651

Stein, G S , Lian, J B , Stein, J L , Van Wijnen, A J , and Montecino, M (1996) Transcriptional control of osteoblast growth and differentiation Physiol Rev 76, 593-629

Stemmer, C , Richalet-Secordel, P , van Bruggen, M , Kramers, K , Berden, J , and Muller, S (1996) Dual reactivity of several monoclonal anti-nucleosome autoantibodies for double-stranded DNA and a short segment of histone H3 J Biol Chem 271 21257-21261

Stephenson, E C , Erba, H P , and Gall, J G (1981) Histone gene clusters of the newt notophthalmus are separated by long tracts of satellite DNA Cell 24, 639-647

Stoler S , Keith, K C , Curnick, K E , and Fitzgerald Hayes, M (1995) A mutation in CSE4, an essential gene encoding a novel chromatin-associated protein in yeast causes chromosome nondisjunction and cell cycle arrest at mitosis Genes Dev 9 573-586

Sullivan K F , Hechenberger, M , and Masri, K (1994) Human CENP-A contains a histone H3 related histone fold domain that is required for targeting to the centromere J Cell Biol 127, 581-592

Sun J M , Wiaderkiewicz, R , and Ruiz-Carrillo, A (1989) Histone H5 in the control of DNA synthesis and cell proliferation Science 245, 68-71

Sun, J M Ali, Z , Lurz, R , and Ruiz-Carrillo, A (1990) Replacement of histone H1 by H5 in vivo does not change the nucleosome repeat length of chromatin but increases its stability EMBO J 9, 1651-1658

Sun J Pilch, D R and Marzluff, W F (1992) The histone mRNA 3' end is required for localization of histone mRNA to polyribosomes, Nucleic Acids Res 20 6057-6066

Swenson, K I , Borgese, N , Pietrini, G , and Ruderman, J V (1987) Three translationally regulated mRNAs are stored in the cytoplasm of clam oocytes Dev Biol 123, 10-16

Szidonya, J , and Reuter, G (1988) Cytogenetic analysis of the echinoid (ed), dumpy (dp) and clot (cl) region in Drosophila melanogaster, Genet Res 51 197-208

Takagaki, Y and Manley, J L (1994) A polyadenylation factor subunit is the human homologue of the Drosophila suppressor of forked protein Nature 372 471-474

Takami, Y , Takeda, S , and Nakayama, T (1995a) Targeted disruption of H2B-V encoding a particular H2B histone variant causes changes in protein patterns on two-dimensional polyacrylamide gel electrophoresis in the DT40 chicken B cell line J Biol Chem 270, 30664-30670

Takami Y , Takeda, S , and Nakayama, T (1995b) Targeted disruption of an H3-IV/H3-V gene pair causes increased expression of the remaining H3 genes in the chicken DT40 cell line J Mol Biol 250, 420-433

Tantravahi, J Alvira, M , and Falck-Pedersen, E (1993) Characterization of the mouse beta maj globin transcription termination region a spacing sequence is required between the poly(A) signal sequence and multiple downstream termination elements Mol Cell Biol 13, 578-587

Tautz, D and Pfeifle, C (1989) A non-radioactive in situ hybridization method for localization of specific RNAs in Drosophila embryos reveals translational control of the segmentation gene hunchback Chromosoma 98, 81-85

Tax, W J M , Kramers, C , van Bruggen, M C J and Berden, J H M (1995) Apoptosis, nucleosomes, and nephritis in systemic lupus erythematosus Kidney Int 48, 666-673

Termaat, R M , Assmann, K J , Dijkman, H B , van Gompel, F , Smeenk, R J and Berden, J H M (1992) Anti-DNA antibodies can bind to the glomerulus via two distinct mechanisms Kidney Int 42, 1363-1371

Thatcher, T H and Gorovsky, M A (1994) Phylogenetic analysis of the core histones H2A, H2B, H3, and H4 Nucleic Acids Res 22, 174-179

Thatcher, T H , MacGaffey, J , Bowen, J , Horowitz, S , Shapiro, D L , and Gorovsky, M A (1994) Independent evolutionary origin of histone H3 3-like variants of animals and Tetrahymena Nucleic Acids Res 22, 180-186

Thoma, F , Koller, T , and Klug, A (1979) Involvement of histone H1 in the organization of the nucleosome and of the salt-dependent superstructures of chromatin J Cell Biol 83, 403-427

Thomas, J O and Rees, C (1983) Exchange of histones H1 and H5 between chromatin fragments A preference of H5 for higher-order structures Eur J Biochem 134, 109-115

Throckmorton, L H (1975) The phylogeny, ecology and geography of Drosophila, In Handbook of genetics Invertebrates of genetic interest Vol 3, King, R C , ed (New York. Plenum Press), pp 421-469

Tonjes, R , Munk, K , and Doenecke, D (1989) Conserved organization of an avian histone gene cluster with inverted duplications of H3 and H4 genes J Mol Evol 28, 200-211

Turner, B M (1993) Decoding the nucleosome Cell 75, 5-8

Turner, P C , Aldridge, T C , Woodland, H R , and Old R W (1983) Nucleotide sequences of H1 histone genes from Xenopus laevis A recently diverged pair of H1 genes and an unusual H1 pseudogene Nucleic Acids Res 11, 4093-4107

Turner, B M and Fellows, G (1989) Specific antibodies reveal ordered and cell-cycle-related use of histone-H4 acetylation sites in mammalian cells Eur J Biochem 179, 131-139

Turner, B M , O Neill, L P , and Allan, I M (1989) Histone H4 acetylation in human cells Frequency of acetylation at different sites defined by immunolabeling with site-specific antibodies FEBS Lett 253, 141-145

Turner, B M , Birley, A J , and Lavender, J (1992) Histone H4 isoforms acetylated at specific lysine residues define individual chromosomes and chromatin domains in Drosophila polytene nuclei Cell 69, 375-384

van Bruggen, M , Kramers, K , Walgreen, B , Elema, J , Kallenberg, C , van den Born J , Smeenk, R , Assman, K , Muller, S , Monestier, M , and Berden, J H M (1996) Nucleosomes and histones are present in glomerular deposits in human lupus nephritis Nephrol Dial Transplant , in press

van Daal, A , White, E M , Gorovsky, M A , and Elgin, S C (1988) Drosophila has a single copy of the gene encoding a highly conserved histone H2A variant of the H2A F/Z type Nucleic Acids Res 16, 7487-7497

van Daal, A , White, E M , Elgin, S C , and Gorovsky, M A (1990) Conservation of intron position indicates separation of major and variant H2As is an early event in the evolution of eukaryotes J Mol Evol 30, 449-455

van Daal, A and Elgin, S C (1992) A histone variant, H2AvD, is essential in Drosophila melanogaster Mol Biol Cell 3, 593-602

van Duin, M , van den Tol, J , Hoeijmakers, J H , Bootsma, D , Rupp, I P , Reynolds, P , Prakash, L and Prakash, S (1989) Conserved pattern of antisense overlapping transcription in the homologous human ERCC-1 and yeast RAD10 DNA repair gene regions Mol Cell Biol 9, 1794-1798

Vanin, E F (1985) Processed pseudogenes characteristics and evolution Ann Rev Genet 19, 253-272

Vasserot, A P , Schaufele, F J and Birnstiel, M L (1989) Conserved terminal hairpin sequences of histone mRNA precursors are not involved in duplex formation with the U7 RNA but act as a target site for a distinct processing factor Proc Natl Acad Sci USA 86, 4345-4349

Vellard M , Sureau, A , Soret, J , Martinerie, C and Perbal, B (1992) A potential splicing factor is encoded by the opposite strand of the trans-spliced c-myb exon Proc Natl Acad Sci USA 89, 2511-2515

Wagner, E G and Simons, R W (1994) Antisense RNA control in bacteria, phages, and plasmids Annu Rev Microbiol 48, 713-742

Wahle, E (1995) 3'-end cleavage and polyadenylation of mRNA precursors Biochim Biophys Acta 1261, 183-194

Wahle, E and Keller, W (1992) The biochemistry of 3'-end cleavage and polyadenylation of messenger RNA precursors Annu Rev Biochem 61, 419-440

Wang, F , Hanske, M , Miedema, K , Klein, G , Ekblom, P , and Hennig, W (1992) Laminin in the male germ cells of Drosophila J Cell Biol 119, 977-988

Waterborg, J H (1990) Sequence analysis of acetylation and methylation in two histone H3 variants of alfalfa J Biol Chem 265, 17157-17161

Waterborg, J H (1991) Multiplicity of histone H3 variants in wheat, barley, rice, and maize Plant Physiol 96, 453-458

Waterborg, J H (1992) Existence of two histone H3 variants in dicotyledonous plants and correlation between their acetylation and plant genome size Plant Mol Biol 18, 181-187

Waterborg, J H (1993) Histone synthesis and turnover in alfalfa Fast loss of highly acetylated replacement histone variant H3 2 J Biol Chem 268, 4912-4917

Waterborg, J H , Winicov, I , and Harrington, R E (1987) Histone variants and acetylated species from the alfalfa plant Medicago sativa Arch Biochem Biophys 256, 167-178

Weiner, A M Deininger, P L and Efstratiadis, A (1986) Nonviral retroposons· genes, pseudogenes, and transposable elements generated by reverse flow of genetic information Ann Rev Biochem 55, 631-661

Wellman, S E , Casano, P J , Pilch, D R , Marzluff, W F , and Sittman, D B (1987) Characterization of mouse H3 3-like histone genes Gene 59, 29-39

Wells, D (1986) Compilation analysis of histones and histone genes Nucleic Acids Res 14 Suppl r119-r149

Wells, D and Kedes, L (1985) Structure of a human histone cDNA evidence that basally expressed histone genes have intervening sequences and encode polyadenylylated mRNAs Proc Natl Acad Sci USA 82, 2834-2838

Wells, D , Hoffman, D , and Kedes, L (1987) Unusual structure, evolutionary conservation of non-coding sequences and numerous pseudogenes characterize the human H3 3 histone multigene family Nucleic Acids Res 15, 2871-2889

Wells, D and McBride, C (1989) A comprehensive compilation and alignment of histones and histone genes Nucleic Acids Res 17 Suppl, 311-346

Wells, D and Bains, W (1991) Characterization of an unusual human histone H3 3 pseudogene DNA Seq 2, 125-127

West, M H and Bonner, W M (1980) Histone 2A, a heteromorphous family of eight protein species Biochemistry 19, 3238-3245

Whitfield, W G , Fellows, G , and Turner, B M (1986) Characterization of monoclonal antibodies to histone 2B Localization of epitopes and analysis of binding to chromatin Eur J Biochem 157, 513-521

Wickens, M (1990) How the messenger got its tail addition of poly(A) in the nucleus Trends Biochem Sci 15, 277-281

Wilhelm, M L and Wilhelm, F X (1989) Histone genes in Physarum polycephalum transcription and analysis of the flanking regions of the two H4 genes J Mol Evol 28, 322-326

Williams, A S , Ingledue, T C , Kay, B K and Marzluff, W F (1994) Changes in the stem-loop at the 3' terminus of histone mRNA affects its nucleocytoplasmic transport and cytoplasmic regulation, Nucleic Acids Res 22, 4660-4666

Williams, A S and Marzluff, W F (1995) The sequence of the stem and flanking sequences at the 3'end of histone mRNA are critical determinants for the binding of the stem-loop binding protein, Nucleic Acids Res 23, 654-662

Witkiewicz, H , Bolander, M E and Edwards, D R (1993) Improved design of riboprobes from pBluescript and related vectors for in situ hybridization Biotechniques 14, 458-463

Wolffe, A P (1989) Dominant and specific repression of Xenopus oocyte 5S RNA genes and satellite I DNA by histone H1 EMBO J 8, 527-537

Wolffe, A P and Pruss, D (1996a) Chromatin Hanging on to histones Curr Biol 6, 234-237

Wolffe, A P and Pruss, D (1996b) Deviant nucleosomes The functional specialization of chromatin Trends Genet 12, 58-62

Wolffe, A P and Pruss, D (1996c) Targeting chromatin disruption Transcription regulators that acetylate histones Cell 84, 817-819

Woodland, H R (1980) Histone synthesis during the development of Xenopus FEBS Lett 121, 1-10

Wu, R S and Bonner, W M (1981) Separation of basal histone synthesis from S-phase histone synthesis in dividing cells Cell 27, 321-330

Wu, R S , Tsai, S , and Bonner, W M (1982) Patterns of histone variant synthesis can distinguish G0 from G1 cells Cell 31, 367-374

Wu, R S , Tsai, S , and Bonner, W M (1983) Changes in histone H3 composition and synthesis pattern during lymphocyte activation Biochemistry 22, 3868-3873

Wu, R S , West, M H P , and Bonner, W M (1984) Patterns of histone gene expression in human and other mammalian cells In Histone genes Structure, organization and regulation, G S Stein, J L Stein, and W F Marzluff, eds (New York John Wiley and Sons), pp 457-483

Wu, R S , Panusz, H T , Hatch, C L , and Bonner, W M (1986) Histones and their modifications CRC Crit Rev Biochem 20, 201-263

Wunsch, A M and Lough, J (1987) Modulation of histone H3 variant synthesis during the myoblast-myotube transition of chicken myogenesis Dev Biol 119, 94-99

Yang, Y S , Brown, D T , Wellman, S E and Sittman, D B (1987) Isolation and characterization of a mouse fully replication-dependent H1 gene within a genomic cluster of core histone genes J Biol Chem 262, 17118-17125

Yuki, S , Inouye, S , Ishimaru, S and Saigo, K (1986) Nucleotide sequence characterization of a Drosophila retrotransposon, 412 Eur J Biochem 158, 403-410

Zlatanova J and Doenecke, D (1994) Histone H1^{0} A major player in cell differentiation? FASEB J 8, 1260-1268

Zlatanova, J and Van Holde, K (1996) The linker histones and chromatin structure New twists In Progress in Nucleic Acid Research and Molecular Biology, W E Cohn and K Moldave, eds (Academic Press Inc), pp 217-259

Zweidler, A (1978) Resolution of histones by polyacrylamide gel electrophoresis in presence of nonionic detergents In Methods in Cell Biology, v XVII Chromatin and Chromosomal protein research II G S Stein, J L Stein and L J. Kleinsmith, eds (New York Academic Press), pp 223-233

Zweidler, A (1984) Core histone variants of the mouse primary structure and differential expression in histone genes In Histone genes Structure, organization and regulation G S Stein, J L Stein, and W F Marzluff, eds (New York John Wiley and Sons), pp 339-371

Summary

The studies presented in this thesis focus on the structure, organization and expression of the replacement histone genes in the fruit fly *Drosophila*. In addition, some novel features of the cell-cycle regulated histone gene expression in *Drosophila* are described.

In Chapter I the literature on the specific properties of the structure, organization, expression and function of the replacement histone genes is reviewed.

In Chapter II the structure, genomic organization and expression of the *Drosophila* genes, encoding the histone H3.3 replacement variant, is described. It is shown that, similar to mammals, there are two non-allelic histone H3.3 encoding genes, H3.3A and H3.3B. Both genes encode exactly the same protein, but they differ in the sequence of their non-coding regions as well as the expression patterns. One of the two genes, H3.3B, produces several different transcripts as a result of alternative polyadenylation.

In Chapter III the properties of the alternatively used polyadenylation sites from the *D. melanogaster* H3.3B gene are analysed in an in vitro polyadenylation system.

In Chapter IV polyadenylated transcripts of the cell-cycle regulated histone genes in *Drosophila* are described. In metazoans, histone genes of this type are normally transcribed into short non-polyadenylated mRNAs with a characteristic stem-loop structure at the 3' end. However, in addition to such mRNAs without a poly(A) tail, cell-cycle regulated histone genes in *Drosophila* generate polyadenylated messengers as well. Firstly, in some cases polyadenylation signals downsteam of the stem-loop are utilized to generate normal polyadenylated transcripts. Secondly, a small proportion of the histone transcripts bear a short poly(A) tail added directly to a partially degraded stem-loop sequence.

In Chapter V a naturally occuring antisense RNA for the cell-cycle regulated histone H3 is described. This antisense RNA is present exclusively in the fly testis, in the nuclei of the primary spermatocytes. This RNA has no coding potential and is likely to have a regulatory function.

In Chapter VI it is shown that the histone H3.3 replacement variant can be distinguished immunologically from its cell-cycle regulated counterparts with the aid of polyclonal antibodies. The H3.3-specific antiserum is used to localize this histone variant in the chromatin of *Drosophila*.

In Chapter VII a novel *Drosophila* histone H4-encoding gene is described. This gene, H4r, encodes a protein, identical to the *Drosophila* cell-cycle regulated histone H4, but has all the properties of a replacement histone gene. This is a first replacement-type histone H4 gene described so far.

In Chapter VIII the sequence of a histone H2B pseudogene is analysed. The structure of this pseudogene suggests that it was generated recently from replication dependent histone H2B gene as a result of a retrotransposition event.

In Chapter IX the results presented in the thesis are discussed in the context of possible future experiments.

Samenvatting

Het onderzoek dat wordt gepresenteerd in dit proefschrift is gericht op de structuur, organisatie en expressie van de 'replacement' histon genen van het fruit vliegje *Drosophila*. In het bijzonder worden enkele nieuwe kenmerken van de cel-cyclus gereguleerde histon gen expressie beschreven.

In hoofdstuk I wordt de literatuur betreffende de specifieke eigenschappen van de structuur, organisatie, expressie en functie van de 'replacement' histon genen samengevat.

In hoofdstuk II wordt de structuur, genomische organisatie en expressie van de *Drosophila* genen, die coderen voor de histon H3.3 'replacement' varianten beschreven. Er wordt laten zien dat, net als bij zoogdieren, er twee niet-allelische histon H3.3 coderende genen zijn, nl. H3.3A en H3.3B. Beide genen coderende voor exact hetzelfde eiwit, maar zij verschillen in de DNA sequentie van hun niet-coderende regio's. Tevens verschillen zij in hun expressie patronen. Het H3.3B gen kan enkele verschillende transcripten synthetiseren op grond van alternatieve polyadenylatie.

In hoofdstuk III worden de eigenschappen van de alternatief gebruikte polyadenylatie regio's van het *D. melanogaster* H3.3B gen nader geanalyseerd m.b.v. een in vitro polyadenylatie systeem.

In hoofdstuk IV worden de polyadenyleerde transcripten van de cel-cyclus gereguleerde histon genen in *Drosophila* beschreven. In metazoa worden mRNA's van histon genen van dit type normaal gesproken gesynthetiseerd zonder poly(A) staart met een karakteristieke 'stem-loop' structuur aan het 3' uiteinde. Hier wordt aangetoond dat de cel-cyclus gereguleerde histon genen van *Drosophila* naast de mRNA's zonder poly(A) staart, ook gepolyadenyleerde transcripten synthetiseren.

In hoofdstuk V wordt een natuurlijk voorkomende antisense RNA van het cel-cyclus gereguleerde histon H3 gen beschreven. Dit antisense RNA komt alleen in de kernen van de primaire spermatocyten van de testis van de vlieg voor. Het heeft geen potentieel om een eiwit te coderen. Het histon H3 antisense RNA heeft waarschijnlijk een regulerende functie.

In hoofdstuk VI wordt laten zien dat de histon H3.3 'replacement' variant immunologisch kan worden onderscheiden van de cel-cyclus gereguleerde typen m.b.v. een polyclonaal antiserum. Het H3.3-specifieke antiserum wordt gebruikt om de H3.3 histon variant in het chromatine van *Drosophila* aan te tonen.

In hoofdstuk VII wordt een nieuw *Drosophila* gen coderende voor histon H4 nader beschreven. Dit gen, H4r, codeert voor een eiwit dat identiek is aan de cel-cylus gereguleerde histon H4 van *Drosophila*. Het H4r gen bezit echter alle eigenschappen van een 'replacement' histon gen. Het is het eerste 'replacement' histon H4 gen tot nog toe beschreven.

In hoofdstuk VIII wordt de sequentie van een histon H2B pseudo-gen nader geanalyseerd. De structuur van dit pseudo-gen suggereert dat het afkomstig is van een cel-cyclus gereguleerde histon H2B gen door retrotranspositie.

In hoofdstuk IX worden de resultaten gepresenteerd in dit proefschrift gedissusieerd in samenhang met mogelijke toekomstige experimenten.

Acknowledgements

This thesis would have never been written without the continuous help and cooperation of many people. In the first place I would like to thank all the colleagues from "Genetica": Cuiying, Hans, Jie, Lisette, Michiel, Petra, Rene, Ron, Wenli, Wolfgang, Xiaoping, Yaxian and Yun. My special thanks are to Harry, for all the help and support, and to Denise, for the excellent work she has done. Miranda, Maja and Wieli are gratefully acknowledged for their technical assistance. I would also like to acknowledge the contribution of Hannie Kremer, whose research formed the starting point of my work.

I am much indebted to Drs. K. Basler, P. Boender, H. Jäckle, E. Hafen, T. Hazelrigg, J. Hoheisel, Ch. Kirchhoff, M. Noll and R. Renkawitz-Pohl for the generous gift of materials. I would like to thank for fruitful cooperation and help Drs. J. Berden, M. van Bruggen and other colleages from the Department of Nephrology (Nijmegen), Drs. M. Polycarpou-Schwarz, S. Gunderson, I. Mattaj (EMBL, Heidelberg), S. Muller (CNRS, Strasbourg) and E. Moudrianakis (Johns Hopkins University, Baltimore).

My greatest thanks are to my loving family. Little Anna, Koos, thank you for sharing with me everything in this thesis and around it.

Curriculum Vitae

Anna Akhmanova was born in Moscow on the 11th of May, 1967. She studied molecular biology in the Moscow State University. After graduation in 1989 she has worked in the A. N. Belozersky Laboratory of the Moscow State University. She has done research on gene stucture and gene regulation in halophilic archaebacteria in the group of Dr. A. S. Mankin.

In 1991 she has moved to the Netherlands, where she was initially a visiting scholar in the Department of Applied Physics in the University of Twente. She was involved in the studies of the interaction of photosensitizing dyes with eukaryotic cells.

In March 1992 Anna Akhmanova has started with her Ph.D. as an "assistent in opleiding" in the Department of Molecular and Developmental Genetics of the Catholic University of Nijmegen under the supervision of Prof. W. Hennig. Research, described in this thesis, was carried out in the period from March 1992 to September 1996. In 1994 she has attended an EMBO Practical Course on the DNA-Transcription Factor Interactions (Naples, Italy) and in 1995 an International Summer School on Post-Transcriptional Control of Eukaryotic Gene Expression (Spetses, Greece).

Since December 1996 Anna Akhmanova is a postdoc in the Department of Microbiology of the Catholic University of Nijmegen.

Anna Akhmanova is married and has one child.