Chapter 25. Nucleosome structure, polarised unwinding and rewinding, and transcriptional regulation.

The access of TFs, transcription- and replication-complexes to open, euchromatic, DNA is under global control during development and maintained in adult tissues. While not being transcribed or replicated, the DNA duplex is wound around nucleosome cores with short linker segments. In addition to the Hox-C functions, chromatin condensation is regulated via additional TATA- and GAGA-binding TFs and chromatin remodelling complexes [1 2 3 4]. The core nucleosome structure consists of about 146 bp of DNA wrapped around four, histone (H) dyads, of between 103 and 136 AAs. The two central H2A/H2B dyads are attached to flanking H3/H4 dyads, which form the DNA entry and exit points [5] Thus, the nucleosome structure consists of an octameric bobbin of chiral histone peptides. It follows that the fine-scale structure of the nucleosome must have only one-fold rotational symmetry, with the DNA strand wound unidirectionally around each histone dyad. An additional "linker" histone, H1, has its central domain inserted along the DNA minor grove, with terminal projections that make contact with adjacent nucleosomes [6 7 8]. The H1 terminal projections may regulate nucleosome packing; while shielding the nucleosome core, and DNA linker segments, from nuclease degradation [9 10]. The nucleosome repeat length (NRL) varies between 176 and 197 bp and is sensitive to the balance between H1 variants and other DNA-binding factors [11 12]. In this context, the residency of the core H2-H4 histones is several hours, with limited access of TFs to buried DNA segments. By contrast, the H1 linker histones have short residency (3-4 minutes) and their loss may expose the linker segments to nuclease degradation [9 13 14 15]. Notably, H1 occupancy is reduced near TSSs and intragenic cis-regulatory sites [16]. During the cell-cycle progression, chromatin compaction is promoted by H1 phosphorylation (via CDK1), while transcription is promoted by H1 dephosphorylation (via CDK2) [9]. Activation of CDK1 may drive the syncytial polar mitotic waves, while the terminal morphogenetic wave in the wing blade is synchronised via CDK1 and the P115 Golgin at the G_2/M checkpoint [17 18 19].

As nucleosomes unwind, the promoter and regulatory segments of the DNA will be exposed. Which of the complimentary DNA strands is transcribed may be dependent on the direction of nucleosome unwinding; which, in turn, reflects the activities of TFs, their cofactors, and the proximity of chromosomal insulator segments. Thus, progressive nucleosome collapse may favour transcription of one, of the two, complimentary DNA strands. Following DNA replication, nucleosomes may re-assemble with alternative phasing patterns, within the active, euchromatic domains. In this context, the H3/H4 entry and exit dimers may be differentially marked. Indeed, alternatively modified (bivalent) chromatin domains have been identified during activation of the zebrafish zygotic transcriptome [20]. Similarly, asymmetric histone marks are present in human embryonic stem-cells and HeLa cultures [21]. Taken together, these data support the hypothesis that nucleosomes may reassemble with alternative phasing patterns in the G_1 and G_2 phases of the cell-cycle. In *Drosophila*, nucleosome stability may be reduced by the En(Z*)* methyl transferase, via the H3 Lysine 27 residue (H3-K27); while Su(var)205 binds methylated H3.3-K27$^{me2/3}$ in centromeres and intercalary heterochromatin [22 23].

As reported in FlyBase, the major *His* gene cluster of *Drosophila* contains multiple copies of the four core histones H2-H4, interspersed with the H1 linker histone. These genes are compact, intron-less and transcribed from alternate DNA strands (>H1<H2B >H2A<H4>H3), without polyadenylation signals (Fig. 32). This organisation should allow rapid transcription of the histone genes in stoichiometric ratios, prior to DNA replication.

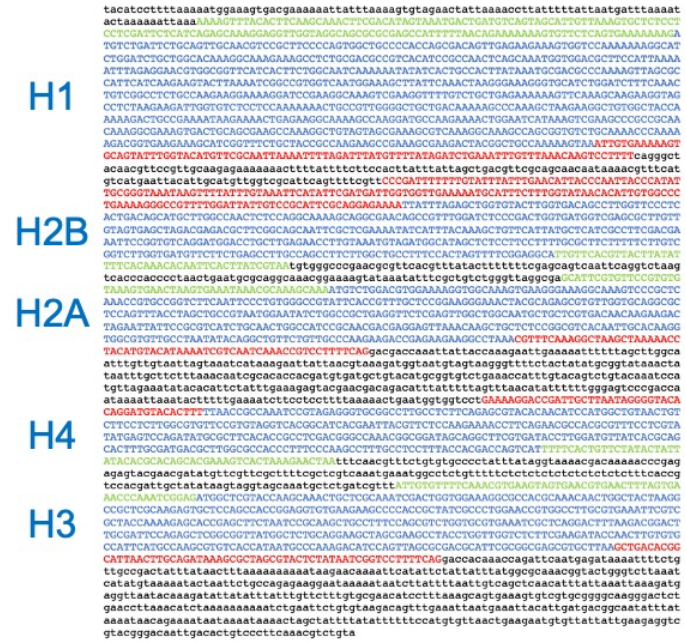

Fig. 32. The major H3 gene cluster of *Drosophila* consistes of a compact cassette, transcribed from alternate DNA strands **[>H1 <H2B >H2A <H4 >H3]**, with the polyA concensus sequence encoded within 3' UTRs. Multiple, tandem repeats of this cassette form the major H3 gene cluster, with strict DNA sequence conservation across coding segments, UTRs and intragenic spacers. However, some of the cassette repeats have truncated H2A genes (lacking 5' or 3' UTRs). 5' UTRs are marked in green typeface, peptide coding sequences in blue, 3' UTRs in red, and intragenic spacers in black (lower case). Modified from FlyBase.

During the rapid syncytial embryonic divisions, the critical bottleneck in nucleosome assembly may be provided by maternally supplied histones, which are gradually released from lipid vesicles. Perhaps surprisingly, the zygotic *His* gene cluster is also transcribed during the acellular blastoderm, and this early zygotic transcription regulates perdurance of the maternally supplied His peptides [24]. During division cycle 14, the major *His* gene cluster forms a discrete histone locus body [24]. Notably, the peptide sequences of the H sub-units are almost completely conserved within the *His* gene cluster, although one of the H1 gene copies (*CG33801*), encodes an additional H^{63} residue. The DNA sequence is conserved across peptide coding sequences, UTRs and intragenic spacers, which presumably reflects strict TF/DNA binding constraints. In consequence, the major *His* gene cluster should allow synchronised transcription, and concerted regulation, of its tandemly repeated gene copies. The number of repeats within the His gene cluster remains uncertain: the FlyBase assembly shows 20 tandem repeats of the cassette, although PCR-based estimates indicate a substantially greater number [25 26 27 28]. The number of *His* gene copies is not critical by the criterion that deletions of the entire cluster are viable, with no apparent haplo-insufficient phenotype. However, no inversion or translocation breakpoints within the His3.3 cluster have been reported, implying that such aberrations may well be lethal. Notably, the major His cluster is flanked by class1 insulator segments, consistent with its regulation as a discrete chromatin domain.

In addition to the major His gene cluster, the *Drosophila* genome encodes several isolated His genes mapping to separate loci: *H3.3A*, *H3.3B*, *BigH1*, *H2AV* and *H4r*.

Nucleosome stability is largely dependent on the balance between the activity of the major *His* gene cluster and these single-copy genes. The single-copy *His* TUs have introns with alternative 5' and 3' UTRs, although each gene encodes only a single peptide isoform. Strikingly, *His3.3A* and *His3.3B* map to separate loci but encode an identical peptide; which differs by three amino acid substitutions (S^{88}>A, V^{90}>I, M^{91}>G) from that encoded by the major His gene cluster. The H3.3 peptide is expressed throughout the cell-cycle and may gradually displace the major H3 Histone in actively transcribed genes. Substitution of the major H3 peptide by H3.3 modifies the silencing of Pc target genes and reduces H1 association with TSSs and regulatory domains [29 16 23]. By contrast, the single *H4 replacement* gene (*H4r*, *CG3379*) encodes a peptide that is identical to that of the major *His* gene cluster. The *H4r* TU has four 5'UTRs, two 3' UTRs, and two introns; consistent with modified regulation of the supplementary H4 in differentiated cell lineages. Within this regulatory network, the BigH1 variant allows rapid nucleosome assembly and maintains zygotic genome silencing before the mid-blastoderm transition [30]. Critically, BigH1 lacks the positively charged AKP helix (Ala, Lys, Pro) encoded by the major H1 gene copies; the AKP helix shields the nucleosome surface and suppresses acetylation of the nucleosome core [7]. By contrast, the H2AV variant has multiple AA substitutions with respect to the major H2A Histone, which it may displace from nucleosomes in response to the SWR1 chromatin remodelling complex. H2AV is required for $H3^{K9}$ methylation and H4K12 acetylation. In this respect, H2AV may be regarded as a PcG function; such that H2AV mutations reduce the proportion of H3-$K9^{me2/3}$ and H4-$K12^{ac}$ marked nucleosomes, and promotes open chromatin [31]. Notably, the $H3^{K27L}$ substitution prevents methylation and blocks Pc target gene repression [26]. In this context, the *H3* variant, *centromere identifier* (*cid*), replaces the major H3 Histone in the kinetochore, and promotes inactive heterochromatin. The Cid histone is incorporated preferentially within the newly synthesised centromere (via Cal1), with a potential function in setting reversed chiral axes in daughter cells (see above, Chapter **22**). Thus, while the major H3.3 gene cluster is expressed in G_1, the single-copy histone genes may drive alterations in nucleosome assembly, stability and collapse throughout the cell-cycle. The chromosomal organisation of the single-copy *his* TUs may allow regulation of alternative nucleosome phasing patterns during the G_1 and G_2 stages of the cell-cycle, within different cell populations.

Summary:

The DNA duplex is wound unidirectionally around octameric histone bobbins, which limit access of nucleases, replicases and TFs. Two central H2A/H2B dyads are flanked by H3/H4 dyads, through which the DNA duplex enters and exits, and an H1 linker histone inserted along the minor DNA groove. In general, the octameric nucleosome bobbin is stable, with core H residences of several hours. However, the linker H1 subunits have shorter residencies of only a few minutes. The major *H* gene cluster in *Drosophila* consists of multiple repeats of a gene cassette without introns or polyadenylation motifs. This chromosomal organisation is consistent with rapid transcription of *H* genes in stoichiometric ratios. In addition, there are several single-copy *H* genes, which have intronic segments and multiple UTRs like other genomic functions. These single-copy H genes may impose global regulation of nucleosome stability in discrete cell lineages, in combination with transcription from the major His3.3 gene cluster. In this context, differentially marked His dyads may assemble into the nucleosome bobbins, with substituted variants, or covalent modifications of the major H peptides.

References:

1. Cairns, B. R. Chromatin remodeling: insights and intrigue from single-molecule studies. *Nat. Struct. Mol. Biol.* **14**, 989–996 (2007).
2. Clapier, C. R. & Cairns, B. R. The biology of chromatin remodeling complexes. *Annu. Rev. Biochem.* **78**, 273–304 (2009).
3. Saunders, A., Core, L. J., Sutcliffe, C., Lis, J. T. & Ashe, H. L. Extensive polymerase pausing during Drosophila axis patterning enables high-level and pliable transcription. *Genes Dev.* **27**, 1146–1158 (2013).
4. Tsai, S.-Y., Chang, Y.-L., Swamy, K. B. S., Chiang, R.-L. & Huang, D.-H. GAGA factor, a positive regulator of global gene expression, modulates transcriptional pausing and organization of upstream nucleosomes. *Epigenetics Chromatin* **9**, 32–32 (2016).
5. Eickbush, T. H. & Moudrianakis, E. N. The histone core complex: an octamer assembled by two sets of protein-protein interactions. *Biochemistry* **17**, 4955–4964 (1978).
6. Luger, K., Mäder, A. W., Richmond, R. K., Sargent, D. F. & Richmond, T. J. Crystal structure of the nucleosome core particle at 2.8 Å resolution. *Nature* **389**, 251–260 (1997).
7. Bharath, M. M. S., Chandra, N. R. & Rao, M. R. S. Molecular modeling of the chromatosome particle. *Nucleic Acids Res.* **31**, 4264–4274 (2003).
8. Syed, S. H. *et al.* Single-base resolution mapping of H1–nucleosome interactions and 3D organization of the nucleosome. *Proc. Natl. Acad. Sci.* **107**, 9620 (2010).
9. Hergeth, S. P. & Schneider, R. The H1 linker histones: multifunctional proteins beyond the nucleosomal core particle. *EMBO Rep.* **16**, 1439–1453 (2015).
10. Zhou, B.-R. & Bai, Y. Chromatin structures condensed by linker histones. *Essays Biochem.* **63**, 75–87 (2019).
11. Routh, A., Sandin, S. & Rhodes, D. Nucleosome repeat length and linker histone stoichiometry determine chromatin fiber structure. *Proc. Natl. Acad. Sci. U. S. A.* **105**, 8872–8877 (2008).
12. Beshnova, D. A., Cherstvy, A. G., Vainshtein, Y. & Teif, V. B. Regulation of the nucleosome repeat length in vivo by the DNA sequence, protein concentrations and long-range interactions. *PLoS Comput. Biol.* **10**, e1003698–e1003698 (2014).
13. van Emmerik, C. L. & van Ingen, H. Unspinning chromatin: Revealing the dynamic nucleosome landscape by NMR. *Prog. Nucl. Magn. Reson. Spectrosc.* **110**, 1–19 (2019).
14. Li, G. & Widom, J. Nucleosomes facilitate their own invasion. *Nat. Struct. Mol. Biol.* **11**, 763–769 (2004).
15. Guertin, M. J. & Lis, J. T. Mechanisms by which transcription factors gain access to target sequence elements in chromatin. *Curr. Opin. Genet. Dev.* **23**, 116–123 (2013).
16. Braunschweig, U., Hogan, G. J., Pagie, L. & van Steensel, B. Histone H1 binding is inhibited by histone variant H3.3. *EMBO J.* **28**, 3635–3645 (2009).
17. Deneke, V. E., Melbinger, A., Vergassola, M. & Di Talia, S. Waves of Cdk1 activity in S phase synchronize the cell cycle in Drosophila Embryos. *Dev. Cell* **38**, 399–412 (2016).
18. Ibar, C. & Glavic, Á. Drosophila p115 is required for Cdk1 activation and G2/M cell cycle transition. *Mech. Dev.* **144**, 191–200 (2017).
19. Vergassola, M, Deneke, V, & Di Talia, S. Mitotic waves in the early embryogenesis of Drosophila: bistability traded for speed. PNAS **115**, E2165–E2174 (2018).
20. Vastenhouw, N. L. *et al.* Chromatin signature of embryonic pluripotency is established during genome activation. *Nature* **464**, 922–926 (2010).

21. Voigt, P. *et al.* Asymmetrically modified nucleosomes. *Cell* **151**, 181–193 (2012).
22. Müller, J. *et al.* Histone methyltransferase activity of a Drosophila Polycomb group repressor complex. *Cell* **111**, 197–208 (2002).
23. Leatham-Jensen, M. *et al.* Lysine 27 of replication-independent histone H3.3 is required for Polycomb target gene silencing but not for gene activation. *PLoS Genet.* **15**, e1007932 (2019).
24. Li, Z. *et al.* Lipid droplets control the maternal histone supply of Drosophila embryos. *Curr. Biol. CB* **22**, 2104–2113 (2012).
25. Lifton, R. P., Goldberg, M. L., Karp, R. W. & Hogness, D. S. The organization of the histone genes in Drosophila melanogaster: functional and evolutionary implications. *Cold Spring Harb. Symp. Quant. Biol.* **42 Pt 2**, 1047–1051 (1978).
26. McKay, D. J. *et al.* Interrogating the function of metazoan Histones using engineered gene clusters. *Dev. Cell* **32**, 373–386 (2015).
27. Bongartz, P. & Schloissnig, S. Deep repeat resolution—the assembly of the Drosophila Histone Complex. *Nucleic Acids Res.* **47**, e18–e18 (2019).
28. Kemp, J. P. J., Yang, X.-C., Dominski, Z., Marzluff, W. F. & Duronio, R. J. Superresolution light microscopy of the Drosophila histone locus body reveals a core-shell organization associated with expression of replication-dependent histone genes. *Mol. Biol. Cell* **32**, 942–955 (2021).
29. Ahmad, K. & Henikoff, S. The histone variant H3.3 marks active chromatin by replication-independent nucleosome assembly. *Mol. Cell* **9**, 1191–1200 (2002).
30. Henn, L. *et al.* Alternative linker histone permits fast paced nuclear divisions in early Drosophila embryo. *Nucleic Acids Res.* **48**, 9007–9018 (2020).
31. Mizuguchi, G. *et al.* ATP-driven exchange of histone H2AZ variant catalyzed by SWR1 chromatin remodeling complex. *Science* **303**, 343–348 (2004).