Chapter 26. Extended TFBS footprint chains, non-linear binding kinetics, and "smeared" genetic enhancer and suppressor functions.

TF target binding sites can be revealed by DNAase1 footprinting [1] [2]. Limited digestion by the DNAase1 endonuclease may remove the H1 linker segments from native chromatin, while leaving the tightly-wound histone cores intact [3]. In general, TF footprints on histone free DNA *in vitro*, should identify core TF/DNA concensus binding sites, with perhaps a few adjacent bases from which the endonuclease might be occluded. However, the length of the protected DNA segment in native chromatin will also depend TF/DNA residency, with the release of short residency TFs allowing progressive degradation, and shortening, of the protected TFBS footprints [4]. Under these conditions, the higher-order binding kinetics of TFs with twin DNA-binding domains may be reflected in their TFBS footprints. Similarly, heteromeric assemblies of multiple TFs should confer longer residency and extended footprint length. In particular, the POZ domain of GAGA-factor (Trl) mediates the formation of heteromeric protein assemblies with increased binding affinity [5]. Thus, the binding of a single TF at an optimum target site may nucleate the assembly of an extended (non-histone) protein sheath. Such heterotypic protein interactions would not be identified by *in vitro* binding assays in the absence of essential co-factors. Under these conditions, the *twi* promoter shows a 40bp DNAase1-protected footprint at a strong Dl concensus binding site, flanked by two weaker Dl sites and four short Zeste (Z) binding sites [6]. This pattern of clustered, discrete TF footprints in the *twi* promoter is consistent with co-operative interactions between DNA-binding proteins *in vivo*.

Similarly, limited DNAase1 digestion of native chromatin can reveal extended TF footprints and nucleosome phasing patterns in synchronised cell cultures. By contrast, only high-residency interactions would be detected in chromatin isolated from asynchronously dividing cell populations. In this context, both the stability of the PollII transcriptional complex and its release may be regulated by TF binding [7] [8] [9]. Thus, a strong TF binding site may initiate the assembly of a heterologous protein sheath, including additional TFs, co-factors and PolII complex components. Once assembled, the transcriptional complex may be paused about 20–50 nucleotides upstream of PolII initiation site in human *myc*, HIV-1 and *Drosophila hsp70* [10] [11] [12]. The few zygotic functions that are transcribed prior to blastoderm cellularisation in *Drosophila*, tend to carry TATA-rich promoters with upstream Zelda sites; while PolII assemblies at GAGA-rich promoters remain paused until their synchronous release during the mid-blastoderm transition [13] [14] [15]. Such paused PolII complexes may set the phase of nucleosome chains extending from 5' promotor segments across UTRs, introns, exons, and intergenic regulatory domains. Notably, putative strong concensus TF binding sites are not restricted to promoter segments, but distributed across protein-coding sequences, introns and intergenic regions, data of [16] analysed via (ALGEN, http://alggen.lsi.upc.es/cgi-bin/promo_v3/promo/promoinit.cgi?dirDB=TF_8.3). Thus, while promotor architecture is critical for transcript initiation, additional regulatory interactions may be smeared across extended chromatin domains. Taken together, these results are consistent with stalled PolII assemblies being nucleated from long residency, cooperative binding sites; with nucleosome phasing patterns set across extended euchromatic domains.

From this perspective, it is notable that zygotic functions transcribed before the mid-blastoderm transition tend to have short transcripts, with few, or no, introns. The initial , transcription of the intronless *WntD* TU (1.14 kb), in the early syncytial blastoderm, is restricted to a few nuclei at the A and P poles, and may supress the nuclear localisation of Dl [17]. However, the *dl* TU (13.5 kb) is transcribed in maternal nurse cells before export to the oocyte, where its mRNA is translated [18]. This maternally supplied Dl perdures during

embryogenesis, while its transcriptional activity is regulated by nuclear/cytoplasmic partitioning via Cact. By contrast, the primary transcripts of the gap genes (*gt*, *hb*, *kr*, *kni*, *tll*, *hkb*, *btd*) are between 2.2 and 7 kb in length, but with extended intragenic regions (13 to 40 kb), which may regulate later developmental functions. In particular, the *hb* TU (6.8 kb) and its surrounding regulatory domains are spanned by a chain of Dl-TFBSs during the mid-blastoderm transition, data of [19]. The individual Dl footprints have a mean length of around 1003 bp, separated by 32-42 bp spacers, consistent with H1 linker segments (Fig.33).

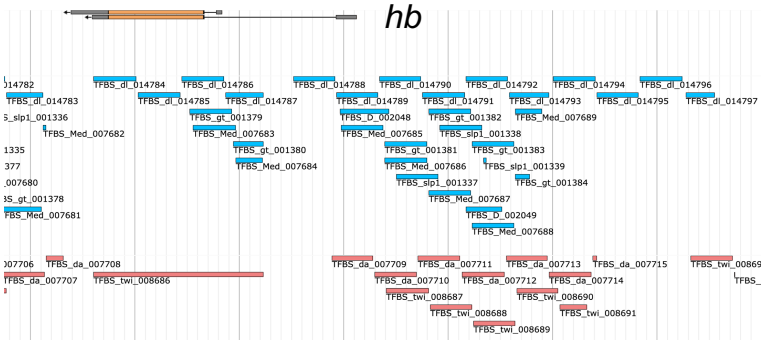


Fig 33. Dl (NF-κβ) footprint chain across *hb* including 5' regulatory region (TFBS dl-14784 to dl-14797). The Helix-Loop-Helix TF footprints, TFBS-Twi and TFBS-Da (orange/red); TFBS Dl, Med, Gt and Slp1 (blue). The TFBS Med, Gt and Slp1 footprint chains are displaced 3' from the TFBS-Dl chain but tend to remain in step. The 32-42 bp spacers between TFBS-Dl footprints are consistent with degradation of H1 linker segments. Data from FlyBase, JBrowse view.

By implication, the Dl-tagged TFBS chains may result from the sequential collapse of six nucleosome (6N) supercoiled stacks. On this hypothesis, the Dl-tagged TFBS chains are consistent with formation of a complex protein sheath covering the unwound DNA duplex. Notably, Dl may act as a transcriptional suppressor, via the mediator complex [20]. Similar, irregular chains of Med-, Gt-, Slp1-, Da- and Twi-tagged footprints are separated by short spacer segments across the *hb* TU; however, Hb-, Cad-, Bcd- and Prd-tagged footprints span irregular, extended domains, without 32-34 bp gaps (https://flybase.org). Presumably these extended "aphasic" footprints correspond to TF-tagged DNA in transcriptionally active segments that are free of nucleosome chains, in different cell populations. The other gap gene TUs show similar out-of-step footprint chains, with differential displacements, although the Dl-TFBS footprints tend to be 5' to Med-TFBS footprints (https://flybase.org). Similar Dl footprint chains extend across the *eve* TU and its intragenic regulatory domains (34).

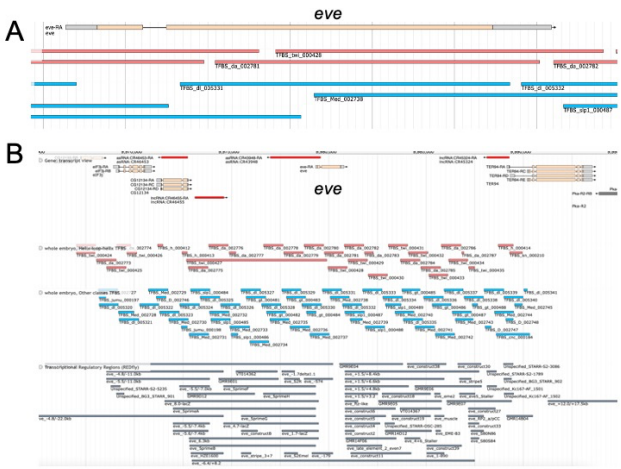


Fig 34. Dl (NF-κβ) footprint chain across *eve*. A. The *eve* gene has a single promotor, a short intron (72 bp) and a single 3' UTR. TFBS Dl, Med, Gt and Slp1 (blue); Twi and Da (orange/red). **B.** A 15 TFBS Dl ladder spans the chromosomal segment surrounding *eve* and adjacent TUs, with 13 Med-, 7 Gt-, 8 Eve- and 4 Ftz-TFBS. Transcriptional regulatory regions (dark grey). from FlyBase, JBrowse view.

Similar TFBS footprint chains span exonic and intronic segments of extended TUs such as *vnd* (16 kb) (Fig. 35).

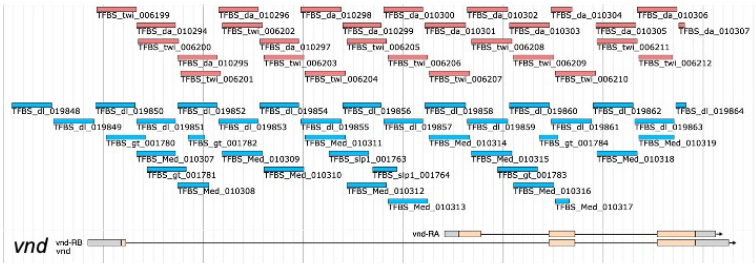


Fig 35. Dl (NF-κβ) footprint chain across *vnd*. TFBS Dl, Med, Gt, Twi and Da tagged footprint chains span both intronic and protein-coding exons of the *vnd* TU (16 kb). TFBS Dl, Med, Gt and Slp1 (blue); Helix-Loop-Helix TFs, Twi and Da, Orange/red. The 3' ends of Med10317, Gt 1784 and Da 10304 may be foreshortened near the splice donor site at the 2nd exon/intron junction. Similarly, Dl-19864 and Da-10307 may be foreshortened before the 3' UTR.

In the case of Trithoraxlike (Trl, GAGA), aphasic, extended footprints may be diced into shorter TFBSs fragments. For example, the *E2F transcription factor-1* (*E2F1*) TU (40.6 kb) is spanned by 24 Trl-tagged TFBSs, most of which match end-to-end without gaps; consistent with precise DNA cuts that remain protected from endonuclease degradation, Fig. 36.

Fig 36. TFBS-Trl fragments in *E2F1*. The E2F1 TU (40.6 kb) is spanned by 24 Trl TFBSs fragments, mostly end-to-end segments, without gaps. By implication, extended Trl-tagged chromatin segments may cover precise DNA cuts that remain protected from DNAaseI degradation. BTB/POZ ChiP TFs (purple), homeodomain TFs (yellow). From FlyBase, JBrowse view.

The *E2F1* gene acts as a transcriptional activator during cell-cycle progression, modulating Cyclin E activity during G_1, and replication functions at the G_1 > S transition [21]. By implication, such Trl-tagged sheaths may splint the cut ends of DNA strands, presumably in cell populations the G_1/S checkpoint. In turn, Cyclin E activity is regulated via the Jak/Stat pathway [22], while an A/P bias in transcription of *E2F1* might be imposed via the Cad gradient [23]. Taken together, these results are consistent with progressive nucleosome collapse of 6N footprint chains initiated from D/V (L/R) midline. Similar TFBS chains extend across many morphogenetic functions, including *vnd*, *msh*, *ind*, *cad*, *dpp*, *ds*, *stan*, *shg*, *numb*, *Myo-II*; the *Antp-C* and *Bx-C*, the *Wnt-C* (*Wnt-4*, *wg*, *Wnt6*, *Wnt10*) the Iroquois-C (*ara*, *caup*, *mir*); and the cognate TUs *inv-en*, *B-H1-B-H2*, *esn-pk*, *cnn-cbs*, *slp1-slp2*, *vk-Col4A*, *knrl-kni* and *dl-Dif*. Some TFBS footprints may be foreshortened near some 5' promoters, splice-acceptor and donor sites, and 3' UTRs. However, although other staggered TFBS chains remain unaffected. This pattern would be consistent with the displacement of low residency chromatin sheaths during PolII progression. Thus, extended TFBS footprint chains may reflect sequential nucleosome collapse following the release of paused transcriptional complexes.

In this context, the Trl (GAGA factor) TF binds methylated H3-K27^{me3} and ubiquinated H2A-K119^{ub1} at PRE sites adjacent to stalled promoters [24 25]. Trl recruits the Polycomb (Pc) group proteins, which negatively regulate Hox TFs and induce chromatin compaction. Meanwhile, SWI/SNF chromatin remodelling complexes remove Pc group proteins, in conjunction with the Zn-finger TFs, Trl and Chinmo. In contrast to Trl, Chinmo modulates chromatin compaction via the Brahma remodelling complex (Actin5C, Actin42A, BAP and Brahma); as well as regulating F-actin assembly within the cytoplasm, via the Jak-Stat pathway [26 27].

In general, histone acetylation promotes an open chromatin configuration, following PolII release from TATA-enriched promoters [9 15 28]. By implication, the transient binding of Hox TFs may set nucleosome phasing patterns, in combination with Zn finger TFs and PolII complex components. During these interactions, Z facilitates recruitment of the Brahma complex to PRE sites [29 30 31]. Notably, the Z protein forms homopolymeric aggregates, which can precipitate *w* DNA *in vitro* (Bickel and Pirrotta, 1990); consistent with the classical genetic analyses of transvection at the *Ubx*, *w*, *y*, *ci* and *dpp* loci [33 34 35 36 37 38 39 40]. Taken together, these studies suggest that Z multimers may couple nucleosome phasing patterns between the maternal and paternal DNA strands. Notably, the maternal and paternal chromatin strands remain in register, as separate, twisted supercoils in polytene chromosomes. In particular, tandem duplications of a large chromosomal segment including the *w* and *rst* (c. 200kb) pair in register, in either cis or trans configurations [36]. Aberration breakpoints that disrupt polytene chromosome pairing also block the z^1-mediated suppression of paired *w* genes in the pigment cells of the eye (Fig. 37).

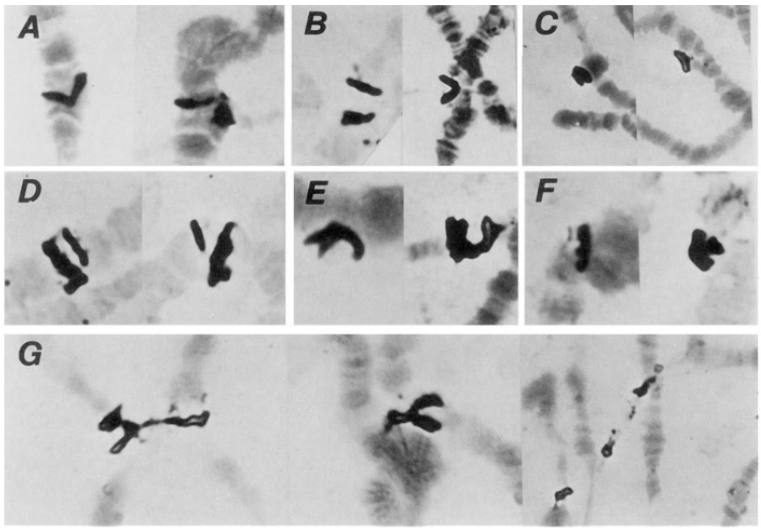
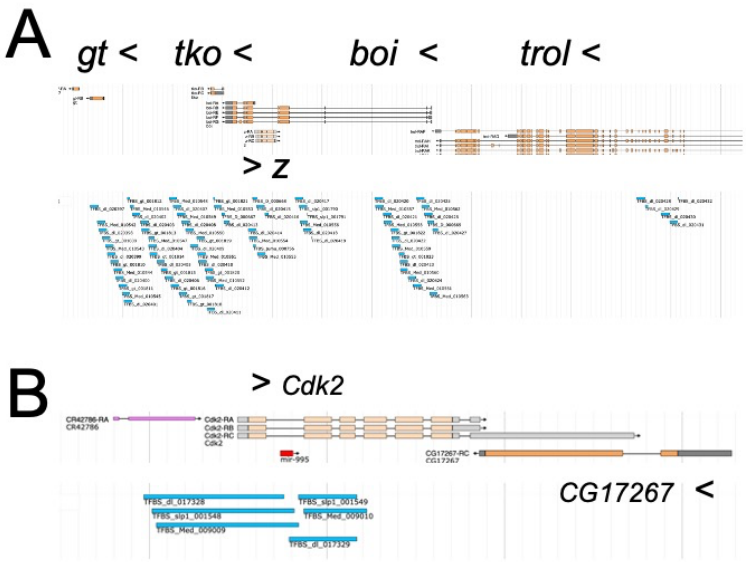


Fig. 37. Polytene chromosomes pair in precise register across the separate maternal and paternal chromatin strands. In TE strains that carry a transposable element with tandem duplications of the *w* - *rst* chromosomal segment, the *w* genes pair in register, in either cis or trans configurations. The maternal and paternal strands remain in separate, twisted supercoils. *In situs* with a *w* probe to salivary gland polytene chromosomes (n = 1024). **A.** Trans-heteozygote between single TE insertion sites displaced by about 60kb, *TE35B(SR100)/TE35BC*. **B.** Heterozygous double-copy TE, with tandem duplication of *w-rst* segment *TE35B(Z)/+* **C.** Reversed double-copy *TE35B(SZ1)/+* in hairpin loop configuration. **D.** Trans-heterozygote between tandem and reversed double-copy TE insertions *TE35B(Z)/TE35B(SZ1)* carrying four copies of the *w-rst* TE. **E.** Triple-copy *TE36B(SR36)/+*, carrying one unpaired *w* gene, which is not suppressed by z^1 in the eye pigment cells. **F.** Three separate views of compact triple-copy *TE36B(SR23)/+*, stretched to different degrees, with one unpaired *w* gene. **G.** *In(2LR) TE35B(SR36)SZ4/+,* a spontaneous inversion with one breakpoint within the triple-copy *TE35B(SR36)SZ4*, allows z^1 suppression of all three *w* genes, which may pair across the inverted segment, *In(2LR)35B;43A*. In all preparations, the polytene chromatin strands stay precise register, despite being wound around separate nucleosome bobbins. From Gubb et al., 1997.

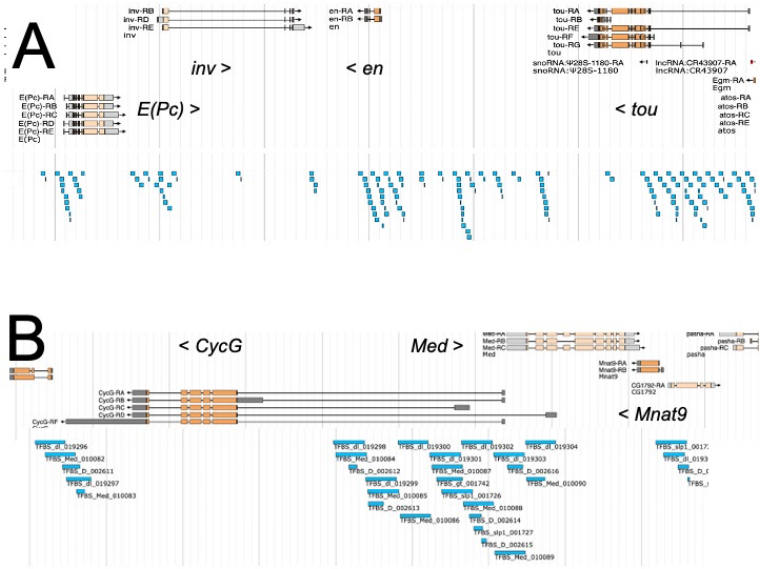
Thus, intermittent coupling by polymeric Z protein aggregates may maintain both diploid and polytene chromatin stands in register. The *z* TU itself is nested within an intron of *boi*, in the opposite orientation to the adjacent morphogenetic functions: < *gt*<*tko*<***b***||>***z***>||<**oi**<*troll*<, (Fig. 38). By implication, *z* transcription may be suppressed during transcription of *boi*, consistent with altered nucleosome phasing during G_1 and G_2.



38. Chromosomal organisation of *zeste*. A. The *z* gene (2.8 kb) is nested within *boi*, in the opposite orientation to the adjacent morphogenetic functions. TFBS footprint chains Dl, Med, and Gt (blue) span the *boi-gt* interval and the 3' UTR of *troll.* **B**. Similarly, the 3' UTR of *Cdk2* overlaps *CG17267*, on the complimentary DNA strand. Note, the Dl-17329, Slp1-1549 and Med-9010 footprints may be foreshortened, consistent with 3' degradation by DNAase1. FlyBase, JBrowse view.

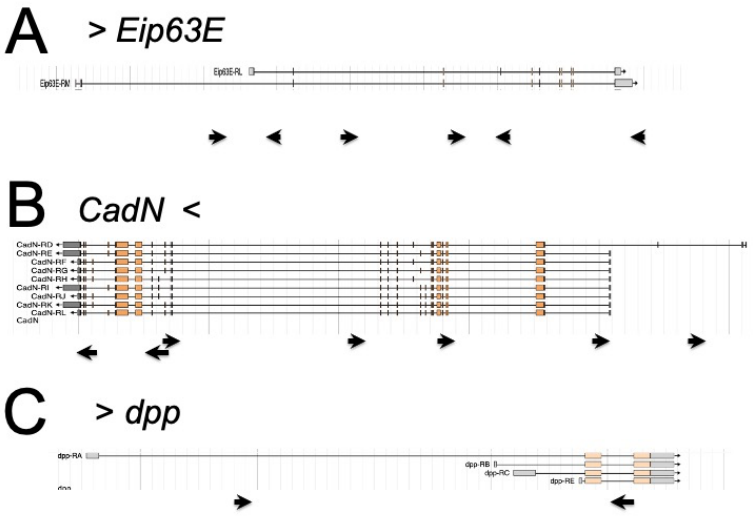
Factors that slow transcription, or stabilise nucleosomes, may act as genetic suppressors; while factors that speed-up transcription, or facilitate nucleosome collapse, may act as genetic enhancers. Whatever the detailed mechanisms, transcription rates may be limited by local activity of TFs and PolII-complex components across intronic segments and extended regulatory domains, in addition to 5' promoter elements and 3' UTRs. In particular, the fluctuating nuclear activities of Dl, Trl and Zelda [41] may alter the differential occupation of TATA-binding sites by the Hox-C TFs (Bcd, Ftz, Zen, Zen2) in competition with Eve, En, Inv and Tbp (TATA-binding protein). The higher-order binding kinetics of heterotypic protein assemblies may be critical for the assembly and release of the PolII transcription complexes. In particular, Tbp is a component of the transcription factor IID complex (TFIID), which is rate-limiting at TATA-box promoters [42]. In this context, the Inv and En TFs might displace each other (and other TATA-binding TFs) from the regulatory domains of downstream morphogenetic functions. Notably, the rare (23) En-tagged footprints are in contrast with the frequent (3, 221) Inv-tagged TFBS footprints, data of [19]. These differential chromosomal distributions are presumably determined by unique DUFs within the Inv and En proteins, together with differential SliM motifs, and the co-factors with which they interact. The short primary transcript of *en* (2.5 kb) is consistent with its early zygotic function; while the extended *inv* TU (35 kb) is expressed only after the post-blastoderm transition [43 44]. Inv-tagged footprints have an average length of about 936 bp, spread across TUs and intragenic regions, without forming footprint chains, data of [19]. By contrast, the En-tagged TFBS sites tend to be adjacent to promoter segments and may be associated with the Exd and Hth Hox-cofactors. In addition, En differs from Inv in having an Eh1 domain (Engrailed homology domain 1). The Eh1 domain mediates Groucho-dependent interactions and is present in many homeodomain (and a few Zn finger) TFs. En may act as a transcriptional suppressor of downstream functions, via Groucho/TLE during the syncytial blastoderm and mid-blastoderm transition. Notably, the *en* promoter carries the PolII stalling motifs: Pause button (PB), Inr,

Nelf-E and GAGA; consistent with its early embryonic functions [45] [46] [25]. By contrast, Inv may act through the SAGA chromatin modifying complex component (Spt3) via CG12112, consistent with acetylation of the H3 histone [47]. The activity of Inv may be modulated by Rel, Akirin, or the NF-κβ factors Dl and Dif [48] [49] [50]. Thus, the partial genetic complementation between the *en*^{1} mutation and *inv* is consistent with their overlapping regulatory domains, rather than reflecting differential DNA-binding affinity of their homoeobox domains, see above Chapter **14**. These regulatory domains may well overlap with adjacent TUs, with *inv* and *en* encoded on complimentary DNA strands, flanked by *enhancer of polycomb E(Pc)* and *toutais* (*tou*), Fig. 39A. The *E(Pc)* > *inv* >|< *en* < *tou* interval includes putative HDAC_PRE sites (Histone Deacetylase Polycomb Repressor Element) and type 1 insulator elements, data of [51] [52] and multiple enhancer segments [53]. The flanking loci, *E(Pc)* and *Tou*, form components of the Histone Acetyl Transferase and ToRC chromatin remodelling complexes, respectively [54] [55]. Thus, this chromosomal organisation is consistent with transcription of one DNA strand increasing nucleosome stability, while transcription the complimentary strand favours nucleosome disassembly. A similar example of cross-regulation between complimentary DNA strands may take place within the *CycG* < > *med* > < *Mnat* 9 interval: with the 3' UTR of *CycG*, and the 5' UTR of *Mnat9*, overlapping the *med* TU, (Fig. 39B).



39. Chromosomal organisation of the *inv en* cognate functions. A. The chromosomal organisation of *inv en*, and adjacent TUs, may reflect cross-regulatory interactions between genes encoded on complimentary DNA strands, with *E(Pc)* and *inv* inverted with respect to *en* and *tou*. TFBS Dl, Med and Slp1 footprints (blue). **B.** Similarly, the 3' UTR of CycG and the 5' UTR of *Mnat9*, overlap the *med* TU in the *CycG* < > *med* > < *Mnat 9* interval. TFBS Dl, Med and D footprints (blue). FlyBase, JBrowse view.

Taken together, these data imply that the regulatory domains of morphogenetic functions may extend across adjacent TUs and be modified by transcription of the complimentary DNA strand. In general, the distribution of *Mi(MIC)* transposon inserts, which can act as splice-acceptors in either orientation, confirms that both DNA strands are transcribed across extended genes, Fig. 40. [53]



40. Active transcription of complimentary DNA strands. *Mi(MIC)* transposons carry a splice-acceptor site attached to -*GFP* These transposon insertions are recovered in either orientation within extended TUs. **A.** *Eip63E* **B.** *CadN* **C.** *dpp.* Black arrows show the orientation of the *Mi(MIC)* insertions. Thus, the recovery of these GFP-expressing insertion strains confirms that both DNA strands are actively transcribed.

On this view of development, transcription of the zygotic genome is dependent on the balance between the activities of genetic enhancers and suppressors. Transcriptional regulation of *dl* takes place in the maternal nurse cells, with imported *dl* mRNA translated in the oocyte. During embryogenesis, the maternal Dl protein perdures through the mid-blastoderm transition, with post-translational control via Cact. Meanwhile, TATA-binding TFs may displace each other from degenerate target sites as the zygotic transcriptome is activated. The post-translational regulation of Dl activity is further modulated by the degradation of Cact by the CalpainA (CalpA) protease; which, in turn, is inhibited by Dpp. In the absence of maternally-supplied Dpp, the polar/equatorial mitotic waves of the syncytial blastoderm are blocked, with ventralisation of the cellular blastoderm [56 57 58]. CalpA also cleaves CycB and is required for the metaphase/anaphase transition [58]. Thus, the nuclear/cytoplasmic partitioning of Dl acts at the core of a complex set of morphogenetic interactions, that are coupled to the cell-cycle progression. Indeed, entangled interactions between the three IκB TFs, *dl*, *dif* (*dorsal-related immunity factor*) and *rel* (*relish*) may regulate many morphogenetic and immune-related functions, with Rel binding to the *IκB* motifs in anti-microbial promotors [59].

By contrast to *en* and *inv*, the TATA-binding *eve* and *ftz* TFs are encoded by widely separated chromosomal loci, with *ftz* embedded within the Ant-C. Both *eve* and *ftz* have short primary transcripts (1.5 and 1.9 kb) with extended intragenic regulatory domains [60] (Small et al. 1992) [62]. In principle, the cofactors that assemble with Eve, or Ftz, may stabilise transient binding to degenerate TATA-binding sites. In practice, the Eve and Ftz cross-regulatory interactions involve a complex set of additional TFs and co-factors [63]. In particular, the Eve repressor activity may regulate parasegmental boundaries via En [64]. On this model, Eve would suppress, and Ftz would enhance, transcription of different set of downstream TUs. The protein/protein interactions (PPIs) of Eve include the Hox TFs: Scr, Antp, Ubx, abd-A, Abd-B; together with the Hox co-factor, Exd; the transcriptional co-repressor, Gro, and the Nmo Kinase. By contrast, the Ftz PPIs include the TATA-binding TFs Ubx and Prd; together with

functions associated with the neuronal and muscle lineages: Ftz-f1 (Ftz transcription factor1), Org-1, Poxn and Chi [65]. Taken together, the striped pattern of Eve/Ftz activities may be set via the early polar transcription of WntD, the nuclear/cytoplasmic partitioning of Dl, and competitive displacement by other TATA-binding TFs. Taken together, these results suggest that TATA binding site occupation may be particularly sensitive to the higher order binding kinetics of individual TFs, together with their cofactors and the (oscillating) nuclear activity of Dl.

Transcriptional regulation may also be affected by the maternal and paternal DNA duplexes being wound around separate histone bobbins. The co-ordinated regulation of diploid genomes must require (at least intermittent) coupling between these separate nucleosome chains. In particular, the Z protein may form polymeric bridging assemblies, with synchronised collapse and re-winding of nucleosome chains. Furthermore, transcriptional release from successive cell-cycle checkpoints may set differential chromatin marks, and incorporate variant His peptides, within assembling nucleosomes. Such alterations may canalise morphogenetic pathways and limit the metabolic range of adult cells.

Summary:

Nucleosome phasing patterns are set during the mid-blastoderm transition as the zygotic transcriptome is fully activated. The progressive collapse of super-coiled nucleosome chains may follow synchronised release of promoter-paused PolII complexes. Thus, promotor architecture drives the polarised collapse of nucleosome bobbins, in conjunction with TFs and TF co-factor activities. In consequence, genetic regulatory domains must co-evolve with the promotors of adjacent TSSs, and chromatin domain boundaries. As the zygotic genome is activated, paused transcription complexes may nucleate the assembly of extended protein sheaths that shield histone-free DNA from endonuclease degradation. Such non-histone sheaths may increase the length of TFBS-tagged footprints and be resistant to exonuclease digestion. Both strong and weak TF concensus-binding sites are distributed across promoters, protein-coding segments, introns and intragenic regions. Weak binding sites may represent degenerate target sites for multiple TFs and confer non-linear binding kinetics, in conjunction with adjacent sites. Thus, the assembly of heterotypic protein sheaths may modify PolII progression, spliceosome assembly and transcript maturation. In general, TFs that impede the transcription of downstream TUs would act as genetic suppressors; while those that facilitate PolII progression would correspond to genetic enhancers. Thus, complex interacting networks of morphogenetic functions may regulate the collapse, and rewinding, of nucleosome bobbins in transcriptionally active domains. In diploid organisms, regulation of the separately wound maternal and paternal nucleosome bobbins may be coupled by multimeric TF concatemers.

References:

1. Galas, D. J. & Schmitz, A. DNAase footprinting a simple method for the detection of protein-DNA binding specificity. *Nucleic Acids Res.* **5**, 3157–3170 (1978).
2. Woodcock, C. L., Skoultchi, A. I. & Fan, Y. Role of linker histone in chromatin structure and function: H1 stoichiometry and nucleosome repeat length. *Chromosome Res. Int. J. Mol. Supramol. Evol. Asp. Chromosome Biol.* **14**, 17–25 (2006).

3. 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).
4. Rusk, N. Transcription factors without footprints. *Nat. Methods* **11**, 988–989 (2014).
5. Espinás, M. L. *et al.* The N-terminal POZ domain of GAGA mediates the formation of oligomers that bind DNA with high affinity and specificity. *J. Biol. Chem.* **274**, 16461–16469 (1999).
6. Pan, D. J., Huang, J. D. & Courey, A. J. Functional analysis of the Drosophila twist promoter reveals a dorsal-binding ventral activator region. *Genes Dev.* **5**, 1892–1901 (1991).
7. Saunders, A., Core, L. J. & Lis, J. T. Breaking barriers to transcription elongation. *Nat. Rev. Mol. Cell Biol.* **7**, 557–567 (2006).
8. Fuda, N. J., Ardehali, M. B. & Lis, J. T. Defining mechanisms that regulate RNA polymerase II transcription in vivo. *Nature* **461**, 186–192 (2009).
9. 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).
10. Bentley, D. L. & Groudine, M. A block to elongation is largely responsible for decreased transcription of c-myc in differentiated HL60 cells. *Nature* **321**, 702–706 (1986).
11. Rougvie, A. E. & Lis, J. T. The RNA polymerase II molecule at the 5′ end of the uninduced hsp70 gene of D. melanogaster is transcriptionally engaged. *Cell* **54**, 795–804 (1988).
12. Gilmour, D. S. & Lis, J. T. RNA polymerase II interacts with the promoter region of the noninduced hsp70 gene in Drosophila melanogaster cells. *Mol. Cell. Biol.* **6**, 3984–3989 (1986).
13. De Renzis, S., Elemento, O., Tavazoie, S. & Wieschaus, E. F. Unmasking activation of the zygotic genome using chromosomal deletions in the Drosophila embryo. *PLoS Biol.* **5**, e117 (2007).
14. Zeitlinger, J. *et al.* RNA polymerase stalling at developmental control genes in the Drosophila melanogaster embryo. *Nat. Genet.* **39**, 1512–1516 (2007).
15. Chen, K. *et al.* A global change in RNA polymerase II pausing during the Drosophila midblastula transition. *eLife* **2**, e00861 (2013).
16. Hoskins, R. A. *et al.* Genome-wide analysis of promoter architecture in Drosophila melanogaster. *Genome Res.* **21**, 182–192 (2011).
17. Ganguly, A., Jiang, J. & Ip, Y. T. Drosophila WntD is a target and an inhibitor of the Dorsal/Twist/Snail network in the gastrulating embryo. *Development* **132**, 3419–3429 (2005).
18. Carroll, S. B., Winslow, G. M., Twombly, V. J. & Scott, M. P. Genes that control dorsoventral polarity affect gene expression along the anteroposterior axis of the Drosophila embryo. *Development* **99**, 327 (1987).
19. Nègre, N. *et al.* A cis-regulatory map of the Drosophila genome. *Nature* **471**, 527–531 (2011).
20. Park, J. M. *et al.* Drosophila Mediator complex is broadly utilized by diverse gene-specific transcription factors at different types of core promoters. *Mol. Cell. Biol.* **21**, 2312–2323 (2001).
21. Crack, D. *et al.* Analysis of Drosophila cyclin EI and II function during development: identification of an inhibitory zone within the morphogenetic furrow of the eye imaginal disc that blocks the function of cyclin EI but not cyclin EII. *Dev. Biol.* **241**, 157–171 (2002).

22. Chen, X. *et al.* Cyclin D-Cdk4 and cyclin E-Cdk2 regulate the Jak/STAT signal transduction pathway in Drosophila. *Dev. Cell* **4**, 179–190 (2003).
23. Hwang, M.-S. *et al.* The caudal homeodomain protein activates Drosophila E2F gene expression. *Nucleic Acids Res.* **30**, 5029–5035 (2002).
24. Espinás, M. L. *et al.* The N-terminal POZ domain of GAGA mediates the formation of oligomers that bind DNA with high affinity and specificity. *J. Biol. Chem.* **274**, 16461–16469 (1999).
25. Hendrix, D. A., Hong, J.-W., Zeitlinger, J., Rokhsar, D. S. & Levine, M. S. Promoter elements associated with RNA Pol II stalling in the Drosophila embryo. *Proc. Natl. Acad. Sci. U. S. A.* **105**, 7762–7767 (2008).
26. Chalkley, G. E. *et al.* The transcriptional coactivator SAYP is a trithorax group signature subunit of the PBAP chromatin remodeling complex. *Mol. Cell. Biol.* **28**, 2920–2929 (2008).
27. Flaherty, M. S. *et al.* chinmo is a functional effector of the JAK/STAT pathway that regulates eye development, tumor formation, and stem cell self-renewal in Drosophila. *Dev. Cell* **18**, 556–568 (2010).
28. Seller, C. A., Cho, C.-Y. & O'Farrell, P. H. Rapid embryonic cell cycles defer the establishment of heterochromatin by Eggless/SetDB1 in Drosophila. *bioRxiv* 450155 (2018).
29. Tamkun, J. W. The role of Brahma and related proteins in transcription and development. *Curr. Opin. Genet. Dev.* **5**, 473–477 (1995).
30. Brown, J. L., Grau, D. J., DeVido, S. K. & Kassis, J. A. An Sp1/KLF binding site is important for the activity of a Polycomb group response element from the Drosophila engrailed gene. *Nucleic Acids Res.* **33**, 5181–5189 (2005).
31. DeVido, S. K., Kwon, D., Brown, J. L. & Kassis, J. A. The role of Polycomb-group response elements in regulation of engrailed transcription in Drosophila. *Development* **135**, 669 (2008).
32. Bickel, Sharon & Pirrotta, V. Bickel S, Pirrotta V. Self-association of the Drosophila zeste protein is responsible for transvection effects. *EMBO J* **9**, 2959–2967 (1990).
33. Gans, M. Etude genetique et physiologique du mutant z de Drosophila melanogaster. *Bull Biol Fr Belg* **38 (Suppl)**, 1–90 (1953).
34. Lewis, E.B. The theory and application of a new method of detecting chromosomal rearrangements in Drosophila melanoga. *Am. Nat.* **88**, 225-239 . (1954).
35. Benson, M. & Pirrotta, V. The product of the Drosophila zeste gene binds to specific DNA sequences in white and Ubx. *EMBO J.* **6**, 1387–1392 (1987).
36. Gubb, D., Roote, J., Trenear, J., Coulson, D. & Ashburner, M. Topological constraints on transvection between white genes within the transposing element TE35B in Drosophila Melanogaster. *Genetics* **146**, 919–937 (1997).
37. Kal, A. J., Mahmoudi, T., Zak, N. B. & Verrijzer, C. P. The Drosophila brahma complex is an essential coactivator for the trithorax group protein zeste. *Genes Dev.* **14**, 1058–1071 (2000).
38. Kassis, J.A. Pairing-sensitive silencing, Polycomb group responseelements, and transposon homing in Drosophila. *Adv Genet.* **46**, 421–438 (2002).
39. Ou, S. A. *et al.* Effects of chromosomal rearrangements on transvection at the yellow gene of Drosophila melanogaster. *Genetics* **183**, 483–496 (2009).
40. Kassis, J. A. Transvection in 2012. Site-specific transgenes reveal a plethora of trans-regulatory effects. *Genetics* **191**, 1037 (2012).
41. Liang, H.-L. *et al.* The zinc-finger protein Zelda is a key activator of the early zygotic genome in Drosophila. *Nature* **456**, 400–403 (2008).

42. Hoffmann, A. *et al.* A histone octamer-like structure within TFIID. *Nature* **380**, 356–359 (1996).
43. Karr, T. L., Ali, Z., Drees, B. & Kornberg, T. The engrailed locus of D. melanogaster provides an essential zygotic function in precellular embryos. *Cell* **43**, 591–601 (1985).
44. Simmonds, A. J., Brook, W. J., Cohen, S. M. & Bell, J. B. Distinguishable functions for engrailed and Invected in anterior–posterior patterning in the Drosopila wing. *Nature* **376**, 424–427 (1995).
45. Goldstein, R. E. *et al.* An eh1-like motif in odd-skipped mediates recruitment of Groucho and repression in vivo. *Mol. Cell. Biol.* **25**, 10711–10720 (2005).
46. Wang, X., Lee, C., Gilmour, D. S. & Gergen, J. P. Transcription elongation controls cell fate specification in the Drosophila embryo. *Genes Dev.* **21**, 1031–1036 (2007).
47. Rhee, D. Y. *et al.* Transcription Factor Networks in Drosophila melanogaster. *Cell Rep.* **8**, 2031–2043 (2014).
48. Nowak, S. J., Aihara, H., Gonzalez, K., Nibu, Y. & Baylies, M. K. Akirin links Twist-regulated transcription with the Brahma chromatin remodeling complex during embryogenesis. *PLOS Genet.* **8**, e1002547 (2012).
49. Bonnay, F. *et al.* Akirin specifies NF-κB selectivity of Drosophila innate immune response via chromatin remodeling. *EMBO J.* **33**, 2349–2362 (2014).
50. Chowdhury, S., Ketcham, S. A., Schroer, T. A. & Lander, G. C. Structural organization of the dynein-dynactin complex bound to microtubules. *Nat. Struct. Mol. Biol.* **22**, 345–347 (2015).
51. Nègre, N. *et al.* A comprehensive map of insulator elements for the Drosophila genome. *PLoS Genet.* **6**, e1000814 (2010).
52. Nègre, N. *et al.* A cis-regulatory map of the Drosophila genome. *Nature* **471**, 527–531 (2011).
53. Cheng, Y. *et al.* Co-regulation of invected and engrailed by a complex array of regulatory sequences in Drosophila. *Dev. Biol.* **395**, 131–143 (2014).
54. Emelyanov, A. V. *et al.* Identification and characterization of ToRC, a novel ISWI-containing ATP-dependent chromatin assembly complex. *Genes Dev.* **26**, 603–614 (2012).
55. Feng, L., Shi, Z., Xie, J., Ma, B. & Chen, X. Enhancer of polycomb maintains germline activity and genome integrity in Drosophila testis. *Cell Death Differ.* **25**, 1486–1502 (2018).
56. Araujo, H. & Bier, E. sog and dpp exert opposing maternal functions to modify Toll signaling and pattern the dorsoventral axis of the Drosophila embryo. *Development* **127**, 3631–3644 (2000).
57. Fontenele, M. *et al.* The Ca2+-dependent protease Calpain A regulates Cactus/IκB levels during Drosophila development in response to maternal Dpp signals. *Mech. Dev.* **126**, 737–751 (2009).
58. Vieira, V., Cardoso, M. A. & Araujo, H. Calpain A controls mitotic synchrony in the Drosophila blastoderm embryo. *Mech. Dev.* **144**, 141–149 (2017).
59. Chowdhury, M. *et al.* An in vitro study of NF-κB factors cooperatively in regulation of Drosophila melanogaster antimicrobial peptide genes. *Dev. Comp. Immunol.* **95**, 50–58 (2019).
60. Frasch, M & Levine, M. Comp1ementary patterns of even-skipped and fushi-tarazu expression involve their differential regulation by a common set of segmentation genes in Drosophila. *Genes Dev.* **1**, 981–995 (1987).
61. Small,S, Blair, A, & Levine, M. Regulation of even-skipped stripe 2 in the Drosophila embryo. *EMBO J* **11**, (1992).

62. Bothma, J. P. *et al.* Dynamic regulation of eve stripe 2 expression reveals transcriptional bursts in living Drosophila embryos. *Proc. Natl. Acad. Sci. U. S. A.* **111**, 10598–10603 (2014).
63. Hughes, S. C. & Krause, H. M. Establishment and maintenance of parasegmental compartments. *Development* **128**, 1109–1118 (2001).
64. Fujioka, M., Yusibova, G. L., Patel, N. H., Brown, S. J. & Jaynes, J. B. The repressor activity of Even-skipped is highly conserved, and is sufficient to activate engrailed and to regulate both the spacing and stability of parasegment boundaries. *Dev. Camb. Engl.* **129**, 4411–4421 (2002).
65. Baëza, M. *et al.* Inhibitory activities of short linear motifs underlie Hox interactome specificity in vivo. *eLife* **4**, e06034 (2015).