Co-localization of Polycomb protein and GAGA factor on regulatory elements responsible for the maintenance of homeotic gene expression

Helen Strutt, Giacomo Cavalli and Renato Paro¹

ZMBH, University of Heidelberg, Im Neuenheimer Feld 282, 69120 Heidelberg, Germany

¹Corresponding author
e-mail: paro@sun0.urz.uni-heidelberg.de

The Polycomb group and trithorax group genes of *Drosophila* are required for maintaining the differential expression state of developmental regulators, such as the homeotic genes, in a stable and heritable manner throughout development. The Polycomb group genes have been suggested to act by regulating higher order chromatin and packaging repressed chromosomal domains in a heterochromatin-like structure. We have mapped, at high resolution, the distribution of Polycomb protein on the bithorax complex of *Drosophila* tissue culture cells, using an improved formaldehyde cross-linking and immunoprecipitation technique. Polycomb protein is not distributed homogeneously on the regulatory regions of the repressed *Ultrabithorax* and *abdominal-A* genes, but is highly enriched at discrete sequence elements, many of which coincide with previously mapped Polycomb group response elements (PREs). Our results further suggest that Polycomb protein spreads locally over a few kilobases of DNA surrounding PREs, perhaps to stabilize silencing complexes. GAGA factor/Trithorax-like, a member of the trithorax group, is also bound at those PREs which contain GAGA consensus-binding sites. Two modes of binding can be distinguished: a high level binding to elements in the regulatory domain of the expressed *Abdominal-B* gene, and a low level of binding to Polycomb-bound PREs in the inactive domains of the bithorax complex. We propose that GAGA factor binds constitutively to regulatory elements in the bithorax complex, which function both as PREs and as trithorax group response elements.

Keywords: bithorax complex/*Drosophila*/GAGA factor/Polycomb/PRE

Introduction

The Polycomb group (PcG) and trithorax group (trxG) genes of *Drosophila melanogaster* are necessary for stabilizing the determined state of expression of homeotic genes during development. They thus maintain the spatial information generated by early patterning mechanisms in the embryo. The trxG proteins are required for transcriptional activity, whereas the PcG proteins are necessary for continued silencing of inactive domains (Paro and Harte, 1996). Thus, in embryos mutant for trxG genes, homeotic

genes initially are expressed normally, but expression fades later in development (Mazo *et al.*, 1990; Breen and Harte, 1993). Conversely, in PcG mutants, target genes become expressed outside of their original boundaries (Struhl and Akam, 1985; Wedeen *et al.*, 1986).

The molecularly characterized members of the trxG were found to consist of a diverse set of proteins. Whilst some proteins, such as the *trithorax* (*trx*) gene product, seem to be specific for regulating developmental genes (Chinwalla *et al.*, 1995), other members appear to have a more general role in gene activation. For example, *brahma* encodes a protein homologous to the yeast SWI2/SNF2 protein (Tamkun *et al.*, 1992) and is a member of a multimeric complex conserved in many species, possessing chromatin reorganization properties (Dingwall *et al.*, 1995; Peterson and Tamkun, 1995). Furthermore, the *Trithorax-like* (*Trl*) gene encodes GAGA factor (Farkas *et al.*, 1994), a sequence-specific DNA-binding protein found in the promoters of many genes, including *Ultrabithorax* (*Ubx*), *engrailed*, *hsp26* and *hsp70* (Biggin and Tjian, 1988; Gilmour *et al.*, 1989; Soeller *et al.*, 1993). Biochemical studies indicate that GAGA factor is able to disrupt nucleosome positioning on the *hsp70* promoter *in vivo* (Tsukiyama *et al.*, 1994). Therefore, it has been suggested that the trxG proteins maintain homeotic gene activity by keeping the DNA in an 'open' chromatin configuration.

The PcG gene products have been shown to associate together in a multimeric complex (Franke *et al.*, 1992). Furthermore, Polycomb protein (PC) and the *Drosophila* heterochromatin protein HP1 share a region of homology (the chromodomain), and therefore it was suggested that the PcG proteins may function by regulating higher order chromatin structure (Paro and Hogness, 1991). This regulation is mediated by PcG response elements (PREs), *cis*-acting DNA elements which are able to maintain the expression boundaries of homeotic genes in a PcG-dependent manner (Müller and Bienz, 1991; Simon *et al.*, 1993). When present in reporter gene constructs in transgenic flies, such elements cause additional PC binding at the insertion site on polytene chromosomes, and are able to silence neighbouring reporter genes (Zink *et al.*, 1991; Chan *et al.*, 1994; Zink and Paro, 1995). PREs are thus thought to act as sequences which recruit members of the PcG and nucleate the formation of PcG silencing complexes on target genes (Pirrotta and Rastelli, 1994; Paro, 1995). Although a number of PREs have been characterized, no clear conserved sequence elements have been identified, and no PcG protein has been demonstrated to bind directly to DNA. Interestingly, immunostaining of polytene chromosomes showed that TRX protein is able to bind to a 14 kb fragment containing the bxd PRE of the BX-C in reporter constructs, suggesting that PREs and trithorax group response elements (TREs) may be closely related (Chinwalla *et al.*, 1995). In addition, transient

transfection assays with reporter constructs encompassing the bxd PRE indicated regulation by both *Pc* and *trx* (Chang *et al.*, 1995).

It was proposed that PcG complexes function by spreading from PREs over extended chromosomal domains. Therefore, as envisaged for heterochromatic silencing, PcG complexes may package chromatin into a compact structure which prevents binding of transcriptional activators to enhancer or promoter sequences (Paro, 1990). This model is supported by the finding that PC is associated with the entire inactive region of the BX-C in tissue culture cells (Orlando and Paro, 1993). However, no difference in accessibility of restriction endonucleases could be detected between active and inactive genes (Schlosser *et al.*, 1994). Similarly, activation by T7 RNA polymerase of a reporter gene inserted in the BX-C was not detectably inhibited (McCall and Bender, 1996), although inhibition of expression of an RNA polymerase II-regulated reporter gene inserted at the same site was observed. It has also been suggested that the PcG silencing complex at PREs has a similar mode of action to that of an activating transcription complex (Bienz and Müller, 1995), and interacts directly with the promoter. Alternatively, the PcG may act by specifically blocking looping interactions between enhancers and promoters (Pirrotta and Rastelli, 1994). Finally, the PcG genes may maintain silencing by sequestering target genes in a particular nuclear compartment (Paro, 1993).

Using an improved methodology for analysing DNA immunoprecipitated from formaldehyde cross-linked chromatin, we were able to analyse PC distribution on the BX-C at a higher resolution. PC is not associated homogeneously with the entire inactive domain, but is highly enriched at discrete sequence elements, which in many cases coincide with characterized PREs. Furthermore, we find that the binding sites of GAGA factor overlap with elements of maximal PC binding.

Results

An improved method for analysing the distribution of proteins on target loci

The distribution of PC on the BX-C was analysed using a formaldehyde cross-linking technique. *Drosophila* Schneider cells were cross-linked *in vivo* and the cross-linked chromatin purified and sheared to an average size of ~1 kb. DNA associated with PC was isolated by immunoprecipitation and amplified by linker-modified PCR, before using as a probe on a Southern of the entire BX-C walk (Orlando and Paro, 1993). This method has now been improved in a number of ways. Firstly, the PCR amplification method has been modified to allow equal amplification of all immunoprecipitated DNA. In the original study, DNA was digested with restriction endonucleases and a cohesive linker attached to create sites for annealing PCR primers. Only DNA fragments with two restriction sites were amplified, and small restriction fragments were amplified much more efficiently than larger fragments, which thus became underrepresented in the final probe. In this report, blunt-end linkers are ligated directly to the immunoprecipitated DNA fragments; such fragments have a random size distribution with respect to the genomic walk and amplify approximately linearly (see

control immunoprecipitation, Figure 1B). Moreover, the PCR amplification step sets the resolution of the technique to ~500 bp, as the final probe consists largely of the smaller sonicated DNA fragments which amplify more efficiently.

Secondly, as control immunoprecipitations without antibody also purify DNA non-specifically, not all signals from PC-immunoprecipitated probes represent true enrichment by PC. To distinguish real immunoprecipitation from background, amplified DNA from both PC and mock immunoprecipitations was immobilized on nylon membranes by slot-blot and hybridized to a variety of probes of the BX-C (data not shown). The resulting signals were quantitated, and the actual enrichment accurately determined. Comparison between a number of fragments analysed by slot-blot allowed the setting of a 'background' level of hybridization in the Southern analysis, as shown by the solid black line in Figure 2. Hybridization signals lower than this level are not considered as being enriched.

Finally, the complete sequence of the BX-C is now available (Martin *et al.*, 1995), thus allowing unambiguous aligning of individual restriction fragments with their position on the walk.

PC protein binds strongly to discrete sequences of the BX-C

DNA purified from PC-immunoprecipitated chromatin was amplified by the blunt end ligation-mediated PCR strategy and used as a probe against the BX-C genomic walk (Figure 1C). All signals in the Southern hybridization were then quantitated by PhosphorImager analysis. The resulting values were normalized to account for molecular weight differences between bands and plotted according to their position on the BX-C walk (Figure 2A).

In agreement with the previous analysis, PC is absent from the coding region of *Abdominal-B* (*Abd-B*) (Figure 2A), which is expressed in Schneider cells (Orlando and Paro, 1993). Note that the background level of hybridization is higher in *Abd-B* than in the rest of the walk. Control experiments indicate that cross-linked chromatin from this region and from other regions containing active genes is sonicated to a relatively smaller average size (data not shown), possibly due to differences in chromatin structure, and is thus expected to amplify more efficiently. In addition, repetitive elements hybridize strongly in both PC and control probes, as previously observed.

Significantly, a discrete number of fragments are highly enriched in PC-immunoprecipitated DNA which are not enriched in the mock immunoprecipitation carried out in parallel (compare Figure 1B and C). Interestingly, many of these peak PC-binding elements correspond to sequences previously identified as having PRE activity. In particular, these peak PC-binding fragments include the Mcp PRE (peak B), the bxd PRE (peak F) and the bx enhancer PRE (peak G), and overlap with the iab3 (peak D) and iab2 (peak E) PREs (Busturia and Bienz, 1993; Qian *et al.*, 1993; Simon *et al.*, 1993; Chan *et al.*, 1994). Therefore, it is likely that the other peak binding sites for PC also correspond to additional, so far uncharacterized PREs. However, PC was not associated with all PRE elements, as poor enrichment was observed at the Fab-7 PRE (Busturia and Bienz, 1993; Zink and Paro, 1995).

Fig. 1. Southern hybridization of immunoprecipitated DNA to the bithorax complex genomic walk. **(A)** One μ g DNA from λ bacteriophage, or 2 μ g of DNA from P1 bacteriophage was digested with restriction enzymes and separated on a 0.6% agarose gel. The λ clones cover the distal region of the BX-C; lane 1 (λ 8106), lane 2 (λ 8099), lane 3 (λ 8095), lane 4 (λ 8088), lane 5 (λ 8083). Lanes 1, 2 and 3 were digested with *EcoRI* and *HindIII*, and lanes 4 and 5 with *EcoRI*. P1 clones cover the remaining 290 kb of the BX-C. Due to the large size (80–100 kb) of insert DNA contained within P1 vectors, many doublet bands were produced after restriction digestion and gel electrophoresis. Therefore, each clone was subjected to two separate digestions, with *EcoRI* or *SalI*–*XhoI*, to allow analysis of all regions of the walk. Distal clones are to the left of the gel, proximal clones to the right; lane 1 (DS04698), lane 2 (DS00846), lane 3 (DS03408), lane 4 (DS03126), lane 5 (DS05563). Lane M is a molecular weight marker, consisting of λ DNA digested with *EcoRI*–*HindIII*; the approximate positions of 5, 4.2 and 2 kb are indicated adjacent to the marker. **(B)** Hybridization of DNA from a control immunoprecipitation (no antibody) to a Southern filter of the gel shown in (A). Note that the hybridization to most restriction fragments is approximately equal, with signal intensity dependent on fragment size. Restriction fragments migrating as doublets show a higher signal intensity, as do those containing repetitive elements; the major repetitive elements are in λ 8088/ λ 8083 (λ clone lanes 4 and 5) and in DS00846 (P1 clone lane 2). The image is identical to that obtained by hybridization of *Drosophila* genomic DNA (data not shown). **(C)** An identical filter to that in (B) hybridized to DNA purified from PC immunoprecipitations. Major elements enriched by PC, with reference to Figure 2A, are a 3.0 kb *HindIII* fragment (peak A) in λ 8106/ λ 8099 (λ clone lanes 1 and 2), 6.0 kb (peak B) and 7.6 kb (peak C) *EcoRI* fragments in DS04698 and DS00846 (P1 clone lanes 1 and 2), 1.9 kb (peak D) and 3.9 kb (peak E) *EcoRI* fragments in DS00846 (P1 clone lane 2), a 3.4 kb *EcoRI* fragment (peak F) in DS03408 (P1 clone lane 3) and a 2.6 kb *EcoRI* fragment (peak G) in DS03126 and DS00563 (P1 clone lanes 4 and 5). **(D)** An identical filter probed with DNA from GAGA factor immunoprecipitations. Note the strong hybridization of a 4.4 kb *EcoRI* doublet in DS04698 (P1 clone lane 1) and a 2.7 kb *EcoRI* fragment in λ 8083/DS04698 (λ clone lane 5 and P1 clone lane 1).

Fig. 2. Distribution of Polycomb and GAGA factor on the bithorax complex. The BX-C is depicted with proximal to the left and distal to the right. The various regulatory regions of the BX-C (abx to iab9) are indicated at the top of the figure. The exon structure of the three homeotic genes is shown in black, and that of other transcripts/ORFs in grey [AHCY, *S*-adenosyl homocysteine hydrolase-like; GLU, glucose transporter-like; W, low density lipoprotein (LDL) receptor-like repeats; X, serine protease-like; Y, chaperonin-containing t-complex protein-1 γ subunit-like; Z, no-on transient A-like; data from Martin *et al.*, 1995]. At the bottom is shown the extent of the P1 or λ clones used in the analysis, together with the restriction sites for *Eco*RI (E), *Sal*I (S) and *Xho*I (X). The traditional map coordinates are shown in bold (Bender *et al.*, 1983; Karch *et al.*, 1985), and in normal type are the coordinates based on the complete sequence of the BX-C (0–340) (Martin *et al.*, 1995). The hybridization signals on the Southern filters in Figure 1 were quantitated, and the resulting values were normalized to account for molecular weight differences and plotted on the map of the BX-C. Data was taken from at least two immunoprecipitation experiments, carried out on independent cross-linked chromatin preparations, and results were similar in each experiment (on average the enrichments varied 10–20% between experiments). Relative PC (A) and GAGA factor (B) binding is shown as grey bars and repetitive elements are shown as white bars with a grey spot below. Restriction fragments showing strong PC binding are labelled A–G on both profiles. The thick black line across each profile represents the approximate background level of hybridization, as determined by slot-blot analysis, and signals below this line are not considered to be enriched. The scale bar on the left indicates enrichment with respect to this background hybridization, which is set at 1.0. The background level in the *Abd-B* locus is higher than in the rest of the walk (dashed black line; see text for details).

Fig. 3. High resolution mapping of Polycomb and GAGA factor at PREs. (A) The seven fragments showing peak levels of PC binding (A–G) were subcloned into the Bluescript KS+ plasmid vector, digested with various restriction enzymes and run on a 1.0% agarose gel. In the cases in which the digests resulted in relatively large fragments, these fragments were purified from agarose gels and further digested. Lane M: molecular weight marker, consisting of λ DNA digested with *EcoRI*–*HindIII*. The approximate positions of 2.0 and 1.0 kb are indicated adjacent to the marker. Peak A: 2982 bp *HindIII* subclone digested with *HindIII*–*AccI* (lane 1) and the two largest fragments further digested; that of 1136 bp with *AseI* (lane 2) and that of 912 bp with *RsaI* (lane 3). Peak B: 5989 bp *EcoRI* subclone digested with *HindIII*–*PstI* (lane 1) and the two largest fragments further digested; that of 2208 bp with *RsaI*–*XmnI* (lane 2) and that of 1556 bp with *PvuII* (lane 3). Peak C: 7652 bp *EcoRI* subclone digested with *BamHI*–*PstI*–*XhoI* (lane 1) and the two largest fragments further digested; that of 1822 bp with *HaeII* (lane 2) and that of 1625 bp with *XmnI* (lane 3). Peak D: a 1890/1894 bp *EcoRI* doublet was resolved by subcloning each into Bluescript. The 1890 bp subclone (lane 1) and the 1894 bp subclone (lane 3) were digested with *EcoRI*; the 1890 bp subclone was further digested with *XhoI*–*Eco47III* (lane 2) and the 1894 bp subclone with *PstI* (lane 4). Peak E: a 3870 bp *EcoRI* subclone digested with *HindIII*–*Sall*–*Clal*–*EcoRI* (lane 1); the largest fragment (1883 bp) further digested with *HaeII* (lane 2). Peak F: a 3384 bp *EcoRI* subclone digested with *KpnI*–*PstI* (lane 1) and the two largest fragments further digested; that of 1210 bp with *Sau3A* (lane 2) and that of 1077 bp with *NdeI* (lane 3). Peak G: a 2571 bp *EcoRI* subclone digested with *HindIII*–*EcoRI*–*Eco47III* (lane 1); the largest fragment (1660 bp) further digested with *Sau3A* (lane 2). (B–D) Southern filters of the gel shown in (A), hybridized to DNA from immunoprecipitations without antibody (B), with PC antibody (C) and with GAGA factor antibody (D).

PC is not associated to a high level with the entire inactive domain of the BX-C. However, DNA adjacent to the peak PC-binding sites is clearly enriched by PC immunoprecipitation. The degree of enrichment appears to decrease with distance from the PRE, indicating that PC associates with a broad domain of 10–15 kb, a possible result of spreading from the peak binding site at PREs.

High resolution mapping of PC-immunoprecipitated DNA

Each peak of PC binding consists of a single restriction fragment. Therefore, we attempted to refine to an even higher resolution the sites at which PC is strongly associated, which could be of importance in PRE function. PC-binding elements were isolated and digested with appropriate enzymes to yield fragments no greater than 1.2 kb (Figure 3A). As before, PC- or control immunoprecipitated DNA was hybridized to Southern filters of

these elements (Figure 3B and C), and the hybridization signals quantitated (Figure 4).

High levels of PC binding in any particular PRE are limited to one or two sub-fragments. In some cases, for example the Mcp PRE (Figure 4B), PC binding is highly restricted, to a single 650 bp DNA fragment. Conversely, binding in the bxd PRE (Figure 4F) is to a larger region of 1.8 kb. Of the known PREs in the BX-C, the bxd element in particular has been well characterized as being a small element (Chan *et al.*, 1994; Chang *et al.*, 1995). Strikingly, our results of PC binding coincide exactly with this minimal PRE element.

Sequence comparison of the PC-immunoprecipitated fragments

No PcG protein has been demonstrated thus far to bind DNA directly. Therefore, it is not clear if PcG proteins are nucleated at PREs by a specific DNA sequence, or

Fig. 4. Distribution of Polycomb and GAGA factor at PREs. Hybridization signals from the filters shown in Figure 3 were quantitated, and normalized to account for molecular weight differences. The total signal for each PRE was used as a guide to normalize the resulting values to the same scale as Figure 2. In each case, the height of the grey bars indicates relative hybridization of DNA immunoprecipitated by PC (left) or GAGA factor (right, where applicable) antibodies, and the scale is the same as that in Figure 2 after subtraction of background hybridization (set at 1.0). The hybridization intensity of the *EcoRI* fragments adjacent to the peak PC-binding elements, as seen on Figure 2, are also shown for comparison. The scale bar in the top right corner of each panel is 1 kb. The vertical black bars above each panel represent consensus GAGA factor-binding sites (minimum sequence GAGAG), and the extents of known PREs are shown as horizontal black bars. (A–G) PC peak binding elements A–G respectively (Figure 2). Restriction enzymes marked are *AccI* (A), *AseI* (As), *BamHI* (B), *ClaI* (C), *EcoRI* (E), *Eco47III* (E4), *HaeII* (Ha), *HindIII* (H), *KpnI* (K), *NdeI* (N), *PstI* (P), *PvuII* (Pv), *RsaI* (R), *Sall* (S), *Sau3A* (S3), *XhoI* (X), *XmnI* (Xm). In each case, proximal is to the left.

alternatively, if PcG complexes recognize a particular DNA conformation at PREs, or interact with other, transient DNA-binding proteins. We compared the sequences of the peak PC-binding elements in the BX-C, and of PC-binding sites in other loci (H.Strutt and R.Paro, unpublished data). Interestingly, strong consensus binding sites for GAGA factor, a trxB protein, were found clustered in many, but not all PREs (black vertical bars above binding profiles in Figure 4). In particular, strong GAGA sites (defined as GA stretches, minimally GAGAG for this analysis) were seen in peak C, peak F (bxd PRE) and peak G (bx PRE), at the exact site of maximal PC binding. Furthermore, GAGA factor-binding sites are seen clustered in peak D in a restriction fragment immediately adjacent to that bound by PC. Peaks A, B (Mcp) and E contain no such clusters of GAGA consensus sequences; however, GAGA factor has been demonstrated to bind some promoter elements lacking a strong consensus, so the absence of such a consensus is not necessarily indicative of a lack of GAGA binding (Granok *et al.*, 1995).

Binding of GAGA factor to consensus sequences in PREs

As GAGA consensus sequences are present at PC-binding elements, we investigated if GAGA protein was associated

with these elements in Schneider cells. Firstly, however, the association of GAGA factor with a known *in vivo* binding site was analysed, to test the efficiency of immunoprecipitation with our GAGA factor antibodies and the accuracy of mapping GAGA factor-binding sites. GAGA factor is constitutively bound to the promoter region of the *hsp26* gene (O'Brien *et al.*, 1995). As expected, chromatin immunoprecipitations specifically enrich restriction fragments from the *hsp26* promoter which contain GAGA consensus-binding sites (Figure 5), but do not enrich for fragments overlapping the transcription unit.

GAGA factor-immunoprecipitated DNA was then used as a probe against a Southern of the BX-C walk (Figures 1D and 2B). Two elements hybridize very strongly with the GAGA factor-immunoprecipitated DNA probe, relative to the mock immunoprecipitation probe. One is a 2672 bp *EcoRI* fragment in the *iab7/iab8* regulatory region of *Abd-B*, the other a 4344/4389 bp *EcoRI* doublet which contains the Fab-7 PRE (Busturia and Bienz, 1993; Zink and Paro, 1995; Hagstrom *et al.*, 1996). Thus the strongest binding of GAGA factor is seen in elements that are not highly enriched for PC, but are in domains required for the regulation of the expressed *Abd-B* gene.

In addition to the strong binding at *Abd-B* regulatory elements, GAGA factor is associated, albeit at a lower

Fig. 5. Binding of GAGA factor to heat shock genes. A map of plasmid pJ1 (Craig, 1980), containing *Drosophila* genomic DNA surrounding the *hsp27*, *hsp23* and *hsp26* heat shock genes, is shown with the restriction sites for *Eco*RI (E) and *Bam*HI (B). Above is an agarose gel containing *Eco*RI–*Bam*HI-digested pJ1 (lane 1), with a marker (M) of λ DNA digested with *Eco*RI–*Hind*III. The approximate positions of 5.0, 2.0 and 1.0 kb are indicated adjacent to the marker. Lanes 2 and 3 are identical Southern blots of this gel, probed with DNA from control or GAGA factor immunoprecipitations respectively. In the control immunoprecipitation, the strongest signal is cross-reaction with the vector (arrow). GAGA factor immunoprecipitations enrich for a 2.8 kb fragment immediately upstream of the *hsp26* transcription start site (the *Eco*RI site is at +6), indicated by the star. DNA surrounding the *hsp27* gene is also enriched, but not DNA covering the *hsp26* transcription unit or the *hsp23* gene.

level, with the four PC-binding elements which contain GAGA consensus sites (peaks C, D, F and G). High resolution mapping confirmed that in every case binding coincides exactly with the presence of clustered GAGA consensus sequences (Figures 3D and 4). The close co-localization of PC- and GAGA factor-binding sites indicates that both proteins can be present on the same regulatory elements of silenced genes.

Although PC and GAGA factor show similar binding at PREs, a dramatic difference can be seen between the distribution of the two proteins. Whereas PC protein shows a broad distribution over a few kilobases surrounding the peak binding site, GAGA immunoprecipitations enrich only those DNA fragments containing GAGA consensus sites, in an all-or-none manner. The difference in the profiles is particularly evident in the high resolution mapping studies (Figure 4): GAGA factor enriches single fragments to a similar degree as does PC, but the neighbouring fragments are not enriched over background. This suggests that PcG complexes form a chromatin structure that is able to spread into DNA adjacent to PREs.

The role of PC and GAGA factor in the regulation of the *Antennapedia* P1 promoter

The finding that both PC and GAGA factor are associated concurrently with identical elements is surprising, as these proteins are expected to have opposing functions. As

PREs and TREs may be closely related (Chang *et al.*, 1995; Chinwalla *et al.*, 1995), one possibility is that GAGA factor is constitutively bound to elements which are also required for *trxG* activation. In addition, the presence of GAGA factor at PC-bound PREs may indicate that GAGA factor also has a function in PcG silencing.

In a first experiment, we examined the genetic interaction between *Pc* and *Trl*. Adult males heterozygous for the alleles *Pc*³ and *Pc*^{XT109} have ectopic sex combs on the second leg. The appearance of these extra sex combs is enhanced in flies which are doubly heterozygous for either *Trl*⁶² or *Trl*^{L3C} (data not shown). This suggests that there may be a role for GAGA factor in PcG silencing, although the effect may not be direct.

Secondly, the function of PC and GAGA factor was examined at an isolated PRE. A 4 kb element of the *Antennapedia* (*Antp*) P1 promoter (from –1.9 to +2.1 relative to the transcription start site) was shown previously to possess PRE activity (Zink *et al.*, 1991). GAGA consensus sequences are seen clustered in two distinct regions. One is a region immediately upstream of the transcriptional start site, consistent with a role for GAGA factor in transcriptional activation at promoters, and the second (containing several strong consensus sequences) is 1.5–1.9 kb upstream from the start of transcription. DNA immunoprecipitated with PC antibodies hybridizes strongly to the upstream element containing GAGA consensus sequences (Figure 6), and more weakly to the promoter element. Strikingly, GAGA factor also binds to the upstream element, but does not bind to the consensus sites near the promoter. This result contrasts with what is observed at the heat shock loci, in which GAGA factor binds constitutively to the promoter region.

To investigate a potential role for GAGA factor in PcG silencing, we analysed transgenic flies containing the *Antp* P1 promoter and PRE upstream of a *lacZ* reporter gene (Zink *et al.*, 1991). In most lines carrying this construct (pAPT 1.0-5C or pAPT 1.8-20B), PC protein is recruited to the insertion site in polytene chromosomes and *lacZ* is not expressed in larval salivary glands. Expression of *lacZ* was examined by β -galactosidase staining of salivary glands in lines heterozygous for the reporter gene construct and the two *Trl* alleles *Trl*^{L3C} and *Trl*⁶². No ectopic β -galactosidase staining was observed in a heterozygous *Trl* background (Table I), indicating that a reduction in GAGA factor concentration does not relieve PcG silencing. It is possible, however, that the reduction of GAGA factor activity was insufficient for an effect to be apparent in this assay. Due to the strong maternal contribution, it is difficult to analyse embryonic effects in the complete absence of GAGA factor (Bhat *et al.*, 1996).

In one line carrying an *Antp* promoter and PRE construct (line pAPT 1.0-79A), PC is not recruited to the insertion site in polytene chromosomes, and the *lacZ* gene is expressed in salivary glands, presumably because of position effects (Zink *et al.*, 1991). *lacZ* expression in salivary glands of this line is much reduced in a *Trl* heterozygous mutant background compared with a wild-type background (Table I), indicating that for the activation function, reduction of *Trl* by one dose results in a measurable effect. This suggests that GAGA factor is required for transcriptional activation, in line with its identity as a *trxG* gene, and that the PRE in the *Antp* P1 promoter also

Fig. 6. Distribution of Polycomb and GAGA factor at the *Antennapedia* P1 promoter PRE. (A) Eight overlapping subclones of the *Antennapedia* promoter region were digested with restriction enzymes to isolate insert DNA from vector, as previously described (Orlando and Paro, 1993) and separated on a 1.2% agarose gel. Only the insert is shown. (B–D) Southern filters of the gel in (A) hybridized with DNA from control (B), PC (C) or GAGA factor immunoprecipitations (D). Both GAGA factor and PC are enriched in clone 1, and PC is also weakly enriched in the promoter region. The lack of strong binding of PC to fragment 1 in the report by Orlando and Paro (1993) was probably due to *Nde*II site distribution. (E) Scheme of the genomic region. A 4 kb *Eco*RI–*Kpn*I fragment (indicated by E and K respectively) surrounding the P1 promoter was analysed. The start of transcription is shown by the arrow, and the positions of the eight subclones in Bluescript are marked.

contains a TRE. The finding that GAGA factor binds to the PRE/TRE in SL2 cells in the absence of transcription may indicate that constitutive binding of GAGA factor is necessary for TRE function.

Discussion

An improved method for analysing protein–DNA interactions *in vivo*

In this report we describe improvements in a formaldehyde cross-linking and immunoprecipitation method, which allow higher resolution mapping of protein–DNA interactions. The PCR amplification step has been modified to allow uniform amplification of all fragments of a genomic walk, and the background level of non-specific immunoprecipitation has been determined more accurately.

The reliability of the new method in the determination of protein-binding sites on DNA *in vivo* has been confirmed by immunoprecipitation experiments against GAGA fac-

Table 1. β -Galactosidase staining of salivary glands of transgenic lines carrying *Antp* P1 promoter fragments in a wild-type or *Trl* background

Transgenic lines	Extent of <i>Antp</i> -P1 promoter fragment	β -gal expression in salivary glands		
		wt	<i>Trl</i> ^{13C}	<i>Trl</i> ⁶²
pAPT 1.8-20B	–9.4 kb to +7.8 kb	–	–	–
pAPT 1.0-5C	–1.9 kb to +2.1 kb	–	–	–
pAPT 1.0-79A	–1.9 kb to +2.1 kb	++	+	+

lacZ expression in *Trl* heterozygous mutants was compared with that of their wild-type siblings.

tor. Whilst PC protein has not been demonstrated to bind directly to DNA, GAGA factor is well characterized in terms of DNA binding, both *in vitro* and *in vivo*. A consensus binding site for GAGA factor, consisting of GAGA repeats, has been proposed. However, GAGA protein can bind to promoter elements lacking such a sequence, and thus any consensus must have additional features (Granok *et al.*, 1995). In fact, GAGA factor immunoprecipitations specifically enrich for DNA containing consensus GAGA-binding sites in the BX-C, *Antp* P1 promoter and the *hsp26* promoter. Furthermore, some GAGA sites, such as the promoter-proximal site of *Antp*, are not enriched, arguing against a general reorganization of chromatin structure during cross-linking allowing artifactual binding to all possible target sequences. Thus, the results produced during this procedure appear to portray accurately the chromatin structure at the time of cross-linking.

PC is highly enriched at PREs in the BX-C

Immunoprecipitations against PC protein strongly and specifically enrich for seven discrete sequence elements in the BX-C. Of these seven elements, five correspond to known PREs, consistent with the idea of PREs being sites for the nucleation of PcG complexes. In particular, the *Ubx* and *abd-A* region of the BX-C have been analysed extensively for elements conveying PRE activity, and it was suggested that there is just one PRE per parasegmental regulatory domain (Chiang *et al.*, 1995). The results shown here confirm this hypothesis in that there is one PC binding peak in each regulatory domain that is inactive in Schneider cells.

The two PREs regulating *Ubx*, in the bxd and bx domains, both correspond to peak PC-binding sites (F and G respectively) in this gene. The bx PRE has only been analysed as a relatively large element, which contains the PC binding peak (Qian *et al.*, 1993; Simon *et al.*, 1993). However, in reporter constructs, a minimal 1.6 kb element has been shown to be sufficient for bxd PRE activity (Chan *et al.*, 1994), and our high resolution mapping indicates that PC binds strongly with the entirety of this element.

We also see PC binding in regions of the iab2, iab3 and iab4 parasegmental domains (peaks E, D and C respectively). In iab2 and iab3, large fragments (11.0 and 11.5 kb respectively) immediately adjacent to the peak, PC sites have been shown to act as PREs in reporter gene constructs (Simon *et al.*, 1993). It has been suggested that PREs may have multiple sub-elements contributing to the final activity (Pirrotta and Rastelli, 1994), which could be

utilized in different spatial or temporal contexts. Therefore, we would expect that elements D and E also have the potential to act as PREs, and that they are the preferred site for PcG nucleation in Schneider cells. The *iab4* regulatory region has not been analysed extensively for elements showing PRE activity. The finding of a PC binding peak in this domain (peak C) suggests that this element is the *iab4* PRE. Finally, PC binds strongly to a discrete 623 bp fragment (peak B) contained within the *Mcp* PRE in *iab5* (Busturia and Bienz, 1993).

No PREs have been identified so far in the more distal region of the BX-C; however, it is likely that peak A, like the other peak PC-binding sites, is an *Abd-B* PRE in *iab9*. Alternatively, it is also possible that peak PC binding is associated with genes in the BX-C unrelated to the homeotic genes. Sequencing of the BX-C revealed the presence of two open reading frames within 10 kb upstream of *Abd-B*; one is a homologue of the human *S*-adenosylhomocysteine hydrolase (*AHCY*) gene, the other shows amino acid identity with a human α -actinin protein (Martin *et al.*, 1995).

No strong PC binding is seen in the regulatory domains *iab6*, *iab7* or *iab8*. In particular, the *Fab-7* PRE, regulating the *iab7* domain, is poorly enriched in PC immunoprecipitations. The *iab5*–*iab9* regulatory regions are each required for modulating differential levels of expression of *Abd-B* in each parasegment of the embryo (Karch *et al.*, 1985). Therefore, in Schneider cells, the lack of PC binding peaks may indicate that the *iab6*–*iab8* domains are positively regulating *Abd-B* expression.

The strong binding of PC at characterized PREs of inactive domains, but not of expressed domains, suggests that the specific association of PC with a PRE is a hallmark of silencing. This would predict an important role for PC in nucleating PcG complexes on the PREs of target genes. Indeed, it was shown previously that tethering a GAL4–PC fusion protein to DNA via artificial GAL4-binding sites was sufficient to recruit an entire PcG complex, and to silence neighbouring genes (Müller, 1995). Conversely, PC is displaced from target PREs when silencing is relieved due to competition from a strong transcriptional activator (Zink and Paro, 1995).

PC protein spreads locally from PREs

It has been suggested that PcG complexes repress target loci by forming complexes of heterochromatin-like structures along the chromosome, thus rendering it inaccessible to transactivating factors (Paro, 1990; Orlando and Paro, 1993). Our results demonstrate that PC does not generate silencing complexes that homogeneously cover the entire regulatory regions of the inactive genes of the BX-C. In fact, PC appears to employ specific interactions with discrete regulatory elements in order to silence target genes, rather than a uniform packaging of chromatin. However, the distribution of PC in the vicinity of PREs suggests that localized spreading of PcG complexes, as opposed to spreading over entire chromosomal domains, may occur to stabilize repression. This apparent spreading is not an artefact due to three-dimensional cross-linking effects, or to poor resolution of the technique, as a similar distribution is not seen with GAGA factor immunoprecipitations. On the contrary, such immunoprecipitations enrich for only those DNA fragments containing GAGA

consensus binding sequences, to a resolution of <1 kb. Neither can the differences in the PC and GAGA factor distributions be attributed to lower sensitivity of the GAGA factor immunoprecipitations, as GAGA factor binding peaks are comparable in intensity with those of PC.

The spreading of PcG complexes is reminiscent of silencing events in the yeast *Saccharomyces cerevisiae*. Telomeres are stably repressed by the SIR2, SIR3 and SIR4 proteins, which are thought to form complexes analogous to PcG complexes (Pillus and Grunstein, 1995). It has been shown that the SIR proteins spread several kilobases from their initial nucleation sites, probably by forming stable interactions with histone proteins (Hecht *et al.*, 1995, 1996). Similarly, limited spreading of PcG complexes, over a few kilobases, may be required for long-term maintenance of silencing.

Spreading of PC is more prominent at some PREs than others (compare the *Mcp* and *bxD* PREs). Thus, DNA flanking the PRE may influence the final distribution of PC. This is in agreement with experiments using a GAL4–PC fusion protein to direct the formation of PcG complexes at artificial GAL4-binding sites, in the absence of a functional PRE. If the fusion protein is removed, the recruited PcG complex is stable throughout many cell divisions if the GAL4-binding sites are flanked by certain *Ubx* sequences, but is not stable if flanked by sequences not normally regulated by the PcG (Müller, 1995). Thus, it was suggested that ‘maintenance elements’, distinct from PREs, are able to stabilize PcG silencing by propagating the spreading of repressive complexes (Paro, 1995).

Alternatively, the spreading we see could be due to the binding of PC to weaker PRE elements in the vicinity of the major PRE. Indeed, it was suggested previously that binding of PC to a major PRE may be stabilized by multiple cooperative interactions with weaker PREs (Pirrotta and Rastelli, 1994). These two possibilities cannot be distinguished from our results, as there is not sufficient resolution to determine if PC contacts the DNA continuously, or is just associated with multiple, closely linked DNA elements.

Although the results with PC seem to argue against long distance spreading as a means of silencing, it is possible that there is a low level of binding over the entire inactive domain, which is below the limits of detection of the method. However, the higher levels of PC at and around PREs indicate that the major role of PC is in the formation of stable complexes at PREs. Other PcG proteins may be distributed more homogeneously in chromatin. It has been predicted that there are 30–40 PcG proteins (Jürgens, 1985), and it would not be surprising if some members of this group, such as PC, are required specifically for nucleating a silencing complex at PREs, whilst other members have more diverse functions.

Binding of GAGA factor to DNA fragments containing PREs

This study shows that in Schneider cells GAGA factor is bound to a number of PREs in the BX-C and to the *Antp* P1 promoter PRE. The strongest GAGA factor binding is in the *Fab-7* element and a more distal 2.7 kb element of the BX-C. Both of these binding sites are located in regulatory domains of the expressed *Abd-B* gene, and are not bound by PC. In addition, a lower level of GAGA

factor binding is seen in PREs in the inactive *Ubx*, *abd-A* and *Antp* genes, at which PC is present. Most characterized GAGA consensus-binding sites have been located in the promoter elements of target genes (Biggin and Tjian, 1988; Gilmour *et al.*, 1989; Soeller *et al.*, 1993). However, it would be consistent with the role of GAGA factor as a trxG gene to find it also associated with more distant regulatory elements. As *Abd-B* is expressed, the strong binding of GAGA factor to two elements in the *Abd-B* regulatory domains is probably related to its activating function, although no GAGA factor binding is seen at the *Abd-B* promoters. Conversely, the lower level of GAGA factor binding may represent a different function of GAGA factor to that of activation, as the *Ubx* and *abd-A* genes are not expressed in Schneider cells, and are bound by PC.

What is the function of GAGA factor at PC-bound PREs? One possibility is that GAGA factor, which is thought to mediate access of *trans*-acting factors to DNA, may also be required to allow access of PcG complexes to target genes. The lack of an effect of *Trl* mutations on an isolated *Antp* PRE does not invalidate such a hypothesis, as GAGA factor activity may not be sufficiently reduced for an effect to be apparent in this assay. In fact, our genetic interaction results suggest that GAGA factor may indeed have a role in PcG silencing.

Our experiments provide support for the idea that GAGA factor is present at PC-bound PREs because these elements also act as TREs. *lacZ* expression in salivary glands from the *Antp* promoter-PRE line pAPT 1.0-79A is 2-fold reduced in a heterozygous *Trl* background, indicating that GAGA factor is a positive regulator of *Antp*. However, as GAGA factor consensus sites are present in the *Antp* promoter region as well as the upstream PRE, it is possible that GAGA factor is also bound to these sites in salivary gland nuclei. In this case, a reduction in promoter activity due to loss of GAGA factor at these sites could also contribute to the reduced *lacZ* expression.

An overlap between TREs and PREs was suggested previously by the finding that TRX and PC proteins are both bound to a number of identical sites on polytene chromosomes, and can also bind simultaneously to a reporter construct containing the bxd PRE in transgenic flies (Chang *et al.*, 1995; Chinwalla *et al.*, 1995). Similarly, GAGA factor is also bound constitutively to the bxd PRE in the absence of transcriptional activation. The finding that GAGA factor binds to many PC-bound PREs suggests that the overlap between PREs and TREs may be widespread.

The level of GAGA factor binding at PC-bound PREs is similar to that at the *hsp26* promoter in the absence of heat shock induction. In this case, GAGA factor is presumed to induce a 'poised' state, to allow rapid gene activation under conditions of stress. The mechanism of trxG activation at more distant elements is less well understood, but it is possible that constitutive binding of GAGA factor and TRX is required for TRE function. Such constitutively bound proteins may be part of a molecular switch, in which the role of GAGA factor is to allow access of other trxG proteins, including additional GAGA proteins, and concomitant loss of PcG proteins when activation occurs, as is observed at Fab-7.

Finally, the binding of GAGA factor to PREs which

Table II. Coordinates of the BX-C subclones

Subclone	BX-C fragment	Coordinates
PC peak A (p8106.1)	2982 bp <i>Hind</i> III fragment	9636–12 618
PC peak B (p5989)	5989 bp <i>Eco</i> RI fragment	109 688–115 677
PC peak C (p7652)	7652 bp <i>Eco</i> RI fragment	123 772–131 424
PC peak D (p1894)	1894 bp <i>Eco</i> RI fragment	152 528–154 422
PC peak E (p3870)	3870 bp <i>Eco</i> RI fragment	159 944–163 814
PC peak F (p3384)	3384 bp <i>Eco</i> RI fragment	218 241–221 625
PC peak G (p2571)	2571 bp <i>Eco</i> RI fragment	273 301–275 872

are silenced by PC is reminiscent of the finding that GAGA factor is bound to other silent DNA domains, the satellite repeats in centromeric heterochromatin (Raff *et al.*, 1994). Mutations in *Trl* have been shown to exhibit defects in the nuclear cleavage cycle, such as failure in chromosome condensation, abnormal chromosome segregation and chromosome fragmentation (Bhat *et al.*, 1996). Furthermore, mitotic defects were also observed in embryos mutant for the heterochromatin protein HP1 (Kellum and Alberts, 1995). It was suggested that GAGA factor is required to organize chromatin structure, in order to allow access of factors mediating condensation and decatenation in domains which would otherwise be tightly packaged (Bhat *et al.*, 1996). Due to the similarity in the mechanisms of PcG and heterochromatin silencing, it is possible that GAGA factor may have a similar role in organizing chromosome structure and segregation at PREs in the BX-C.

Materials and methods

DNA stocks and plasmid subclones

For analysis of genomic walks, DNA was isolated from five bacteriophage P1 clones covering 280 kb of the BX-C (Martin *et al.*, 1995). The P1 clone spanning the remaining 60 kb grew very poorly, and DNA from five bacteriophage λ clones was used to cover the *Abd-B* domain (Karch *et al.*, 1985). For analysis of the *Antp* P1 promoter, eight overlapping subclones of a 4 kb *Eco*RI–*Kpn*I fragment surrounding the promoter were used, as previously described (Orlando and Paro, 1993).

Fragments showing peak Polycomb binding were subcloned into the corresponding restriction enzyme site in the polylinker of Bluescript KS+ for further analysis. The coordinates of the BX-C subclones (Table II) are based on the complete published sequence (SEQ89E; DDBJ/EMBL/GenBank accession No. U31961) of the bithorax complex (Martin *et al.*, 1995).

Fly stocks

For the genetic interactions each of the two *Pc* alleles, *Pc*³ and *Pc*^{XT109} were crossed to two *Trl* alleles *Trl*^{13C} and *Trl*⁶² (Farkas *et al.*, 1994).

Antibodies

Antibodies recognizing bacterially expressed GAGA factor were kindly provided by Peter Becker (EMBL, Heidelberg). Use of these antibodies in Western blot analysis of nuclear extract from embryos or SL2 cells or in staining of *Drosophila* polytene chromosomes gave results similar to those observed with previously published antibodies (Raff *et al.*, 1994; Tsukiyama *et al.*, 1994).

In vivo formaldehyde cross-linking of Schneider cells, and immunoprecipitation of cross-linked chromatin

Cross-linking of *Drosophila* Schneider cells and immunoprecipitation from cross-linked chromatin were carried out exactly as previously described (Orlando and Paro, 1993; Orlando *et al.*, 1997), except for a modification in the PCR amplification protocol, in which a blunt-ended linker was ligated directly to the sonicated, immunoprecipitated DNA. This method amplifies all regions of genomic DNA approximately linearly. However, the average size of DNA resulting from the PCR

amplification is 400–500 bp, considerably smaller than that of the input DNA. This is not surprising, as small DNA fragments amplify more efficiently. This selective amplification indicates that the PCR amplification step, not the initial chromatin size, sets the resolution of the technique.

After purification of immunoprecipitated DNA, approximately one-third of the resulting DNA (1 ng) was resuspended in 7 µl of dH₂O, to which 1 µl of 10× ligation buffer (0.5 M Tris–HCl pH 7.6, 125 mM MgCl₂, 250 mM dithiothreitol, 12.5 mM ATP) was added, and linkers to a final concentration of 0.1 µM. Linkers were made by annealing a 24mer oligonucleotide of sequence 5'-AGA AGC TTG AAT TCG AGC AGT CAG and a 20mer of sequence 5'-CTG CTC GAA TTC AAG CTT CT, of which only the 24mer is phosphorylated at the 5' end. Ligation was carried out at 4°C overnight, after the addition of 4 U of T4 DNA ligase. Amplification was carried out directly in a 100 µl reaction using *Taq* polymerase (Boehringer) and the corresponding buffer, deoxynucleotides at 250 µM, and 1 µM primer (the 20mer oligonucleotide used above). The amplification cycles were 1 cycle of 94°C for 2 min; 35 cycles of 94°C for 1 min, 55°C for 1 min, 72°C for 3 min; and 1 cycle of 94°C for 1 min, 55°C for 1 min, 72°C for 10 min.

Amplified DNA was phenol–chloroform extracted and ethanol precipitated, and linker DNA sequences removed by *Hind*III digestion. Amplified DNA subsequently was purified from linker DNA using Qiagen PCR purification columns, according to the manufacturer's conditions. One hundred ng of purified DNA from plus and minus antibody immunoprecipitations was analysed by slot-blot, or 50–100 ng used as a hybridization probe on Southern blots, as previously described (Orlando and Paro, 1993; Orlando *et al.*, 1997).

Analysis of β -gal expression in salivary glands

Three transgenic *Drosophila* lines containing the *Antp* P1 promoter and PRE upstream of a *lacZ* reporter gene were crossed with lines carrying the *Trl*^{13C} or the *Trl*⁶² mutant alleles over a TM3 balancer chromosome. *lacZ* expression was examined in isolated salivary glands of the F1 third instar larvae by X-gal staining, as previously described (Zink *et al.*, 1991). Half of the salivary glands from the progeny of pAPT 1.0-79A crosses stained ~2-fold less intensively than salivary glands carrying the transposon in a wild-type background. This population of salivary glands was inferred to carry the *Trl* mutation.

Acknowledgements

We thank Peter Becker for GAGA factor antibodies, Ursi Weber and the Berkeley *Drosophila* genome project for BX-C P1 clones, and Gunter Reuter for *Trl* fly stocks. We also thank Valerio Orlando for technical advice and comments on the manuscript. G.C. is supported by an EMBO long term fellowship, and the work of R.P. is funded by the Deutsche Forschungsgemeinschaft and by the Fond der Chemischen Industrie.

References

- Bender, W., Akam, M.E., Karch, F., Beachy, P.A., Peifer, M., Spierer, P., Lewis, E.B. and Hogness, D.S. (1983) Molecular genetics of the bithorax complex in *Drosophila melanogaster*. *Science*, **221**, 23–29.
- Bhat, K.M., Farkas, G., Karch, F., Gyurkovics, H., Gausz, J. and Schedl, P. (1996) The GAGA factor is required in the early *Drosophila* embryo not only for transcriptional regulation but also for nuclear division. *Development*, **122**, 1113–1124.
- Bienz, M. and Müller, J. (1995) Transcriptional silencing of homeotic genes in *Drosophila*. *BioEssays*, **17**, 775–784.
- Biggin, M.D. and Tjian, R. (1988) Transcription factors that activate the *Ultrabithorax* promoter in developmentally staged extracts. *Cell*, **53**, 699–711.
- Breen, T.R. and Harte, P.J. (1993) *trithorax* regulates multiple homeotic genes in the bithorax and Antennapedia complexes and exerts different tissue-specific, parasegment-specific and promoter-specific effects on each. *Development*, **117**, 119–134.
- Busturia, A. and Bienz, M. (1993) Silencers in *Abdominal-B*, a homeotic *Drosophila* gene. *EMBO J.*, **12**, 1415–1425.
- Chan, C.-S., Rastelli, L. and Pirota, V. (1994) A *Polycomb* response element in the *Ubx* gene that determines an epigenetically inherited state of repression. *EMBO J.*, **13**, 2553–2564.
- Chang, Y.-L., King, B.O., O'Connor, M.O., Mazo, A. and Huang, D.-H. (1995) Functional reconstruction of *trans* regulation of the *Ultrabithorax* promoter by the products of two antagonistic genes, *trithorax* and *Polycomb*. *Mol. Cell. Biol.*, **15**, 6601–6612.
- Chiang, A., O'Connor, M.B., Paro, R., Simon, J. and Bender, W. (1995) Discrete Polycomb-binding sites in each parasegmental domain of the bithorax complex. *Development*, **121**, 1681–1689.
- Chinwalla, V., Jane, E.P. and Harte, P.J. (1995) The *Drosophila trithorax* protein binds to specific chromosomal sites and is co-localized with *Polycomb* at many sites. *EMBO J.*, **14**, 2056–2065.
- Craig, E.A. (1980) Four *Drosophila* heat shock genes at 67B: characterization of recombinant plasmids. *Nucleic Acids Res.*, **8**, 4441–4457.
- Dingwall, A.K., Beek, S.J., McCallum, C.M., Tamkun, J.W., Kalpana, G.V., Goff, S.P. and Scott, M.P. (1995) The *Drosophila* snr1 and brm proteins are related to the yeast SWI/SNF proteins and are components of a large protein complex. *Mol. Biol. Cell*, **6**, 777–791.
- Farkas, G., Gausz, J., Galloni, M., Reuter, G., Gyurkovics, H. and Karch, F. (1994) The *Trithorax-like* gene encodes the *Drosophila* GAGA factor. *Nature*, **371**, 806–808.
- Franke, A., DeCamillis, M., Zink, D., Cheng, N., Brock, H.W. and Paro, R. (1992) *Polycomb* and *polyhomeotic* are constituents of a multimeric protein complex in chromatin of *Drosophila melanogaster*. *EMBO J.*, **11**, 2941–2950.
- Gilmour, D.S., Thomas, G.H. and Elgin, S.C.R. (1989) *Drosophila* nuclear proteins bind to regions of alternating C and T residues in gene promoters. *Science*, **245**, 1487–1490.
- Granok, H., Leibovitch, B.A., Shaffer, C.D. and Elgin, S.C.R. (1995) GAGA over GAGA factor. *Curr. Biol.*, **5**, 238–241.
- Hagstrom, K., Müller, M. and Schedl, P. (1996) *Fab-7* functions as a chromatin domain boundary to ensure proper segment specification by the *Drosophila* bithorax complex. *Genes Dev.*, **10**, 3202–3215.
- Hecht, A., LaRoche, T., Strahl-Bolsinger, S., Gasser, S.M. and Grunstein, M. (1995) Histone H3 and H4 N-termini interact with SIR3 and SIR4 proteins: a molecular model for the formation of heterochromatin in yeast. *Cell*, **80**, 583–592.
- Hecht, A., Strahl-Bolsinger, S. and Grunstein, M. (1996) Spreading of transcriptional repressor SIR3 from telomeric heterochromatin. *Nature*, **383**, 92–96.
- Jürgens, G. (1985) A group of genes controlling the spatial expression of the bithorax complex in *Drosophila*. *Nature*, **316**, 153–155.
- Karch, F., Weiffenbach, B., Peifer, M., Bender, W., Duncan, I., Celniker, S., Crosby, M. and Lewis, E.B. (1985) The abdominal region of the bithorax complex. *Cell*, **43**, 81–96.
- Kellum, R. and Alberts, B. (1995) Heterochromatin protein 1 is required for correct chromosome segregation in *Drosophila* embryos. *J. Cell Sci.*, **108**, 1419–1431.
- Martin, C.H., Mayeda, C.A., Davis, C.A., Ericsson, C.L., Knafels, J.D., Mathog, D.R., Celniker, S.E., Lewis, E.B. and Palazzolo, M.J. (1995) Complete sequence of the bithorax complex of *Drosophila melanogaster*. *Proc. Natl Acad. Sci. USA*, **92**, 8398–8402.
- Mazo, A.M., Huang, D.-H., Mozer, B.A. and Dawid, I.B. (1990) The *trithorax* gene, a trans-acting regulator of the bithorax complex in *Drosophila*, encodes a protein with zinc-binding domains. *Proc. Natl Acad. Sci. USA*, **87**, 2112–2116.
- McCall, K. and Bender, W. (1996) Probes for chromatin accessibility in the *Drosophila* bithorax complex respond differently to *Polycomb*-mediated repression. *EMBO J.*, **15**, 569–580.
- Müller, J. (1995) Transcriptional silencing by the Polycomb protein in *Drosophila* embryos. *EMBO J.*, **14**, 1209–1220.
- Müller, J. and Bienz, M. (1991) Long range repression conferring boundaries of *Ultrabithorax* expression in the *Drosophila* embryo. *EMBO J.*, **10**, 3147–3155.
- O'Brien, T., Wilkins, R.C., Giardina, C. and Lis, J.T. (1995) Distribution of GAGA protein on *Drosophila* genes *in vivo*. *Genes Dev.*, **9**, 1098–1110.
- Orlando, V. and Paro, R. (1993) Mapping Polycomb-repressed domains in the bithorax complex using *in vivo* formaldehyde cross-linked chromatin. *Cell*, **75**, 1187–1198.
- Orlando, V., Strutt, H. and Paro, R. (1997) Analysis of chromatin structure by *in vivo* formaldehyde cross-linking. *Methods*, **11**, 205–214.
- Paro, R. (1990) Imprinting a determined state into the chromatin of *Drosophila*. *Trends Genet.*, **6**, 416–421.
- Paro, R. (1993) Mechanisms of heritable gene repression during development of *Drosophila*. *Curr. Opin. Cell Biol.*, **5**, 999–1005.
- Paro, R. (1995) Propagating memory of transcriptional states. *Trends Genet.*, **11**, 295–297.
- Paro, R. and Harte, P.J. (1996) The role of Polycomb group and trithorax group chromatin complexes in the maintenance of determined cell states. In Russo, E. *et al.* (eds), *Epigenetic Mechanisms of Gene Regulation*. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY, pp. 507–528.

- Paro,R. and Hogness,D.S. (1991) The Polycomb protein shares a homologous domain with a heterochromatin-associated protein of *Drosophila*. *Proc. Natl Acad. Sci. USA*, **88**, 263–267.
- Peterson,C.L. and Tamkun,J.W. (1995) The SWI/SNF complex: a chromatin remodeling machine? *Trends Biochem. Sci.*, **20**, 143–146.
- Pillus,L. and Grunstein,M. (1995) Chromatin structure and epigenetic regulation in yeast. In Elgin,S.C.R. (ed.), *Chromatin Structure and Gene Expression*, Oxford University Press, Oxford, pp. 123–146.
- Pirrotta,V. and Rastelli,L. (1994) *white* gene expression, repressive chromatin domains and homeotic gene regulation in *Drosophila*. *BioEssays*, **16**, 549–556.
- Qian,S., Capovilla,M. and Pirrotta,V. (1993) Molecular mechanisms of pattern formation by the BRE enhancer of the *Ubx* gene. *EMBO J.*, **12**, 3865–3877.
- Raff,J.W., Kellum,R. and Alberts,B. (1994) The *Drosophila* GAGA transcription factor is associated with specific regions of heterochromatin throughout the cell cycle. *EMBO J.*, **13**, 5977–5983.
- Schlosser,J., Eggert,H., Paro,R., Cremer,S. and Jack,R.S. (1994) Gene inactivation in *Drosophila* mediated by the *Polycomb* gene product or by position effect variegation does not involve major changes in the accessibility of the chromatin fibre. *Mol. Gen. Genet.*, **243**, 453–462.
- Simon,J., Chiang,A., Bender,W., Shimell,M.J. and O'Conner,M. (1993) Elements of the *Drosophila* bithorax complex that mediate repression by *Polycomb* group products. *Dev. Biol.*, **158**, 131–144.
- Soeller,W.C., Oh,W.E. and Kornberg,T.B. (1993) Isolation of cDNAs encoding the *Drosophila* GAGA transcription factor. *Mol. Cell. Biol.*, **13**, 7961–7970.
- Struhl,G. and Akam,M. (1985) Altered distributions of *Ultrabithorax* transcripts in *extra sex combs* mutant embryos of *Drosophila*. *EMBO J.*, **4**, 3259–3264.
- Tamkun,J.W., Deuring,R., Scott,M.P., Kissinger,M., Pattatucci,A.M., Kaufmann,T.C. and Kennison,J.A. (1992) *brahma*: a regulator of *Drosophila* homeotic genes structurally related to the yeast transcriptional activator SNF2/SWI2. *Cell*, **68**, 561–572.
- Tsukiyama,T., Becker,P. and Wu,C. (1994) ATP-dependent nucleosome disruption at a heat-shock promoter mediated by binding of GAGA transcription factor. *Nature*, **367**, 525–532.
- Wedeer,C., Harding,K. and Levine,M. (1986) Spatial regulation of *Antennapedia* and bithorax gene expression by the *Polycomb* locus. *Cell*, **44**, 739–748.
- Zink,B., Engström,Y., Gehring,W.J. and Paro,R. (1991) Direct interaction of the *Polycomb* protein with *Antennapedia* regulatory sequences in polytene chromosomes of *Drosophila melanogaster*. *EMBO J.*, **10**, 153–162.
- Zink,D. and Paro,R. (1995) *Drosophila* *Polycomb*-group regulated chromatin inhibits the accessibility of a *trans*-activator to its target DNA. *EMBO J.*, **14**, 5660–5671.

Received on February 6, 1997; revised on March 14, 1997