Reference: Cavalli, G., orlando, V and Paro, R. (1999). Mapping DNA target sites of chromatin-associated proteins by formaldehyde cross-linking in *Drosophila* embryos. pp20-37. in Chromosome Structural Analysis: A Practical Approach. Ed. Bickmore, W.A. Oxford University Press.

Mapping DNA target sites of chromatin-associated proteins by formaldehyde cross-linking in *Drosophila* embryos

Giacomo CAVALLI, Valerio ORLANDO and Renato PARO

1. Introduction

The method described here allows the mapping of protein-DNA interactions through the ability of formaldehyde to cross-link proteins and nucleic acids in living cells. Formaldehyde is a very reactive dipolar compound which reacts with the amino groups of proteins and amino acids (1, 2). It shows no reactivity however towards free double-stranded DNA, and thus does not cause the extensive DNA damage seen after prolonged exposure to other cross-linking reagents such as UV. Each formaldehyde molecule has the capacity to interact with two amino groups. Therefore DNA-protein, protein-protein and RNA-protein cross-links are rapidly formed after formaldehyde treatment, creating a stable structure which prevents the redistribution of cellular components. Furthermore, a simple heat treatment is sufficient to reverse the reaction equilibrium, and to allow isolation of pure DNA for further analysis (3). Formaldehyde cross-linking combined with chromatin immunoprecipitation (IP) is a way of mapping the *in vivo* distribution of chromatin-associated proteins. As such, this technique is of great value in the analysis of protein-DNA interactions, even more so when studying proteins which do not show specific DNA binding activities *in vitro*. This has recently been demonstrated for several chromatin-associated proteins, such as Polycomb (Pc) in *Drosophila* (4) and the SIR proteins in budding yeast (5). Additionally, analysis of the cross-linking pattern can not only allow mapping of sites of protein-DNA interaction, but can also give an estimation of the relative binding affinity to different sequences across a large genomic region.

The method presented here is based on a previously described method of formaldehyde cross-linking in *Drosophila* cultured Schneider SL-2 cells (4). This method was recently improved with a modification of the PCR amplification step in order to allow a more accurate quantification of relative binding affinities for adjacent sequences with a resolution in the order of 500 bp (6). Here we present the adaptation of this methodology to the analysis of chromatin from *Drosophila* embryos. Although SL-2 cells have been successfully used as a model system for the study of various cellular processes, the analysis of embryonic chromatin permits additional functional studies. For example, the competition between different proteins for binding to overlapping target sites during development might be investigated at endogenous sites or in transgenic constructs.

Moreover, the wealth of genetic mutants of *Drosophila* allows study of the DNA binding of multimeric protein complexes in the absence of single components.

In this chapter, we will discuss the different steps involved in the mapping of *in vivo* DNA binding sites by formaldehyde cross-linking, focusing in particular on the steps of preparation and cross-linking of the embryos. As an example of the application of the method, we will show the analysis of the binding profile of two proteins. The first is Pc, a protein involved in the maintenance of the repressed state of homeotic genes from mid embryogenesis to the adult state (7). The second is GAGA factor, which counteracts Pc-mediated silencing maintaining the spatially restricted activation pattern of homeotic genes (8).

2. Outline of the method

The analysis of *in vivo* DNA protein interactions by formaldehyde cross-linking involves the following steps:

- Preparation of embryos for cross-linking
- Formaldehyde cross-linking and purification of soluble cross-linked chromatin
- IP of purified cross-linked chromatin, reversal of the cross-links and DNA purification
- PCR-amplification of the immunoprecipitated DNA
- Analysis of the immunoprecipitated DNA, which can be done in two ways:
  (i) by analysis of the enrichment of specific DNA fragments in slot-blots with probes which may represent putative target sites of the protein under study.
  (ii) by using PCR-amplified DNA as a radioactively labeled probe to hybridise a Southern blot of DNA from chromosome walk of a genomic region of interest

3. Formaldehyde cross-linking in staged *Drosophila* embryos.

3.1 Preparation of fly cages and collection of staged embryos

For analysis of cross-linked chromatin of 11-16 h (after egg lay) old embryos, *Drosophila* embryo collection was performed at 25°C in cylindrical fly cages of 30 cm diameter and 35 cm depth containing 7500 to 15000 flies per cage. Flies were 5-12 days old. Two fly cages were used for wild type (Oregon R) flies, while four cages were necessary for the transgenic line 5F24 25,2, which lays significantly less eggs then the wild type flies. Six apple-juice agar plates of 14 cm diameter, distributed on two levels in the cage were used for collection. Under these conditions about 1 g wet weight of 11-16 h old embryos could be collected (in our experience, a minimum of 0.5 g is required in order to obtain enough DNA for PCR amplification). For staging, a pre-lay collection of 2 h with plates streaked with 45% acetic acid and fresh yeast paste was first made. Collection was from 5 p.m. to 10 p.m. Collection plates were incubated overnight at 25°C. At 9 a.m., embryos

were collected in embryo wash buffer and further processed as described in *Protocol 1*. Essentially, embryos were dechorionated, washed and cross-linked. Cross-linked embryos were extensively washed and then sonicated in order to produce soluble chromatin, which was used for further analysis (see below).

Many chromatin proteins might change their target size distribution at different developmental stages, therefore a careful staging is essential in these experiments. In order to monitor staging, nuclei from a small aliquot of the cross-linked embryos were stained with the DNA staining dye Hoechst 33258. Stained embryos were analysed by fluorescence- and light microscopy in order to asses the average stage of development. We found that preparations containing a large proportion of embryos older then 16 h resulted in a high fraction of uncross-linked chromatin, most likely due to the poor permeability of these embryos to formaldehyde. On the other hand, cross-linking of embryos during early development can be easily performed, but it must be noted that the same weight equivalent contains a smaller number of nuclei, i.e. of chromatin. This is particularly important when studying early stages, such as pre-blastoderm embryos, where up to 5-10 g wet weight might be required to obtain DNA amounts sufficient to represent the whole genome complexity after IP.

With this setup, up to 4 embryo collections could be obtained from wild type fly cages in alternate days. On the days where no cross-linking was performed, flies were fed in the morning and in the evening with three apple-juice agar plates/cage, on which about 20 g of fly medium and 2 g of yeast paste were streaked. When transgenic flies were used, only 2-3 embryo collections per cage could be performed.

3.2. Optimizing cross-linking conditions

Permeabilization of the embryos was obtained by adding three volumes of n-heptane directly to the cross-linking solution (*Protocol 1*), and cross-linking with vigorous shaking. This was sufficient to permeabilize embryos until late developmental stages, but it seems not sufficient for cross-linking of embryos at the very end of embryogenesis or upon hatching, as judged by the high loss of material which was observed from preparations containing a high fraction of 16-20 h old embryos. These embryos were also refractory to Hoechst staining. Thus, more severe methods of fixation should be tested for these late stages.

The concentration of formaldehyde required for optimal cross-linking has been tested extensively using different concentrations of formaldehyde. The optimal range was found to be between 1.8% and 3.7%. The effect of undercross-linking is shown in *Figure 1*. Using 1% formaldehyde, a large fraction of the embryonic chromatin is not cross-linked, resulting in the fractionation of a large amount of free DNA to the bottom of the gradient (*Figure 1B, 1D*). Sometimes, this was also observed using 1.4% formaldehyde. On the other hand, a concentration of 1.8% resulted in a reproducible cross-linking of about 80% of the material, eluting in a fraction with the density characteristic of protein-DNA complexes (δ=1.39, see *Figure 1A, 1C*). Concentrations of 2.8% and of 3.7% gave a pattern indistinguishable from that observed with 1.8% formaldehyde, while higher concentrations resulted in loss of a large fraction of the material and to DNA of a high molecular weight (M.W.), probably since highly cross-linked material is refractory to shearing by sonication.

Protocol 1. Formaldehyde cross-linking of *Drosophila* embryos

Equipments and reagents

- Sonifier apparatus (Branson Ultrasonics Corporation, Sonifier Model 250) equipped with a Microtip 3/16" (Cat. N. 101-148-069).
- 0.1 mm diameter glass beads
- 3% NaOCl.
- Cross-linking solution: 1.8% formaldehyde, 50 mM HEPES, 1 mM EDTA, 0.5 mM EGTA, 100 mM NaCl, pH 8.0. Add formaldehyde immediately before use, from a 37 % stock solution stabilised with 10 % methanol.
- n-heptane.
- Glycine.
- PBS.
- 100% Glycerol.
- Hoechst 33258 (Sigma).
- Embryo wash buffer (EWB): 0.03% TritonX100, 0.4% NaCl.
- Wash solution A: 10 mM HEPES pH 7.6, 10 mM EDTA pH 8.0, 0.5 mM EGTA pH 8.0, 0.25% Triton X100.
- Wash solution B: 10 mM HEPES pH 7.6, 200 mM NaCl, 1 mM EDTA pH 8.0, 0.5 mM EGTA pH 8.0, 0.01% Triton X100.
- Sonication buffer: 10 mM HEPES pH 7.6, 1 mM EDTA pH 8.0, 0.5 mM EGTA, pH 8.0.
- 10 % N-lauroylsarcosine
- CsCl, optical grade.

Method

1. Dechorionate approximately 1 g embryos in 3% NaOCl in EWB for 2-3 min at room temperature (r.t.). Wash extensively with EWB. Transfer the embryos to a 50 ml falcon tube and wash once in 0.01% Triton X100 made up in PBS.

2. Cross-link with 10 ml of cross-linking solution for 15 min in the presence of 30 ml n-heptane. Shake vigorously.

3. Spin embryos for 1 min in a tabletop centrifuge at maximum speed. Stop cross-linking by washing with 50 ml of PBS, 0.125 M glycine, 0.01% Triton X100. Let embryos sediment without centrifuging.

4. DNA stain a small aliquot of embryos to check embryo staging as follows: wash embryos with 0.5 ml PBS. Add 250 µl 1µg/ml Hoechst 33258 in PBS, stain 4 min in the dark, let embryos sediment. Wash twice with 0.4 ml PBS. Resuspend embryos in 25 µl PBS (at this stage embryos can be stored overnight at 4°C in the dark). Add 25 µl of 100% glycerol and mix well. Mount on a slide and coverslip and analyse on fluorescence microscope with a suitable filter. Score 100-200 embryos to control staging.

5. Wash embryos in 15 ml (falcon tubes) with wash solution A for 10 min on a roller. Repeat with wash solution B. Resuspend in 5.5 ml (final volume) sonication buffer. Up to this stage, embryos should still be intact. This can be checked under a microscope.

6. Sonicate with Branson Model 250 sonifier equipped with a microtip. Sonication has to be empirically adjusted and the sample must be kept on ice throughout. We recommend immersion of the tip about 2 cm deep in the solution. Perform four sonication cycles of 30 sec. at constant power in the presence of 0.1 mm glass beads. Power must be gradually increased up the maximum level possible, avoiding foaming. It should be possible to reach level 6 to 8 in the output control scale of the Branson Model 250 sonifier. Pause 90 sec. between each cycle. At the end, inspect a 10 µl aliquot under phase contrast microscopy to ensure that all nuclei are lysed (no large particles should be left).

7. Adjust to 0.5% N-lauroylsarcosine and rotate for 10 min. Spin debris at high speed (in microfuge tubes for 5 min). The supernatant can be further processed by CsCl gradient purification (*Protocol 2*) immediately, or frozen in liquid N_2 and stored for several days before further purification. This is particularly useful when working with fly lines that lay poorly, so that several cross-linking preparations can be pooled together.

Protocol 2**. Purification of soluble cross-linked chromatin**

Equipment and reagents

- Dialysis buffer: 4 % glycerol, 10 mM Tris-HCl pH 8.0, 1 mM EDTA pH 8.0, 0.5 mM EGTA pH 8.0.
- Dialysis bags: micro-collodion bags (Sartorius, Catalog N. 13202).
- 1 mg/ml RNase A (DNase-free)
- 10 mg/ml Proteinase K (in TE buffer, stored at -20°C).
- 3 M sodium acetate (NaOAc) pH 5.2.
- 100% Ethanol (on ice)

Method

1. To supernatant from *Protocol 1* add sonication buffer (*Protocol 1*) with 0.5% N-lauroylsarcosine to a volume of 8 ml. Add 5.68 g CsCl (i.e. adjust to a density of 1.42g/ml) and make up to 10 ml final volume with sonication buffer with 0.5% N-lauroylsarcosine.

2. Divide the sample between two 5 ml tubes to avoid overloading the gradient (after centrifugation a broad sarkosyl/lipid/protein aggregate is present at the top of the gradient, and overloading can result in poor resolution of the gradient, and thus poor quality chromatin). Spin CsCl gradient for 72h, 20°C, 195000 *g* (40000 rpm in Beckman rotor SW55Ti).

3. Elute 12 x 400 µl fractions per gradient with a peristaltic pump at about 1 ml/min with tubings of 0.114 cm (0.045 inches) internal diameter.

4. Check the density profile of fractions (should be in range of 1.390 to 1.362 refraction index)

5. Dialyse fractions in micro-collodion bags against dialysis buffer. After 2 h change the buffer and continue dialysis overnight.

6. After the buffer change in step 5., remove about 1/10 vol (40 µl) to a microfuge tube to check the size and quantity of DNA in each gradient fraction (see *Figure 1*) as follows: Add RNase A to 50 µg/ml and incubate 30 min at 37°C. Add Proteinase K to 500 µg/ml and SDS to 1%. Incubate at 56°C for 1 h to partially reverse the cross-links. Purify DNA once by phenol/chloroform extraction, ethanol precipitate and resuspend in 15 µl 1X loading buffer for agarose gel electrophoresis. Check DNA samples on a 1% agarose gel.

7. Pool appropriate dialysed chromatin fractions together (usually good fractions have densities of between 1.350 and 1.450 g/ml and, depending on efficiency of sonication, DNA is from 0.2 kb to 20 kb in size with an average size of 1.0 kb). Freeze 500 µl aliquots in liquid N_2 and store at -80°C. These aliquots should contain 30-60 µg DNA. To analyse DNA of the pooled fractions, take 50 µl of the final pool and repeat RNase treatment, Proteinase K/SDS, and DNA purification of step 6. Estimate DNA concentration by measuring the optical density at 260 nm and run an agarose gel to check DNA size-distribution. Frozen chromatin aliquots are suitable for IP and should be stable for several months.

4. Immunoprecipitation of cross-linked embryonic chromatin and PCR amplification of the immunoprecipitated DNA

IP of the cross-linked purified chromatin, reversal of the cross-links and DNA purification are performed as previously described (4, 9) and detailed in *Protocol 3*. Recently, an improvement of the PCR amplification step has been introduced in order to allow a more reliable quantification of the relative enrichment of neighbouring DNA fragments (6). The purified DNA is directly ligated to a blunt ended linker and PCR-amplified. This assumes that (i) chromatin DNA fragments which are not blunt-ended are not ligated to the linker, and (ii) among the ligated molecules, short fragments, in the range of 200-500 bp, are amplified more efficiently than longer ones. Thus, a fraction of the cross-linked DNA molecules is lost during the procedure. Since this loss is random throughout the genome, it does not compromise the quantification as long as the whole genomic complexity is represented in the samples before PCR amplification. Amplification artifacts may be observed under non-optimal sonication conditions, since too large a proportion of very long fragments results in a correspondingly high loss of DNA molecules during PCR. This might result in loss of entire regions of genomic DNA in the PCR-amplified samples. Such artifacts can be identified by analysis of the hybridisation profile of a PCR-amplified, radioactively labeled, sample from a mock IP to a restriction digest of a large genomic walk. Since no antibody is used in the mock IP, the probe obtained from this sample should hybridise uniformly to all fragments of the digest in the absence of artefactual loss of genomic complexity (see *Protocol 4* and also *Figure 3A, B, and C*, mock IP samples). Finally, it should be noted that the preferential amplification of shorter DNA fragments sets the resolution of the technique to about 500 bp.

Protocol 3. Chromatin immunoprecipitation

Reagents

- Protein A Sepharose CL4B (PAS, Sigma). Equilibrate in RIPA buffer (below) by mixing at 4°C for 30 - 60 min. 100 mg PAS equilibrated in 1 ml RIPA buffer results in a 50 % v/v suspension. After equilibration, PAS is stable for up to 1 week at 4 °C.
- RIPA buffer: 140 mM NaCl, 10 mM Tris-HCl pH 8.0, 1 mM EDTA, 1 % Triton X-100, 0.1 % SDS, 0.1 % sodium deoxycholate, 1 mM PMSF (on ice). Add PMSF immediately before use from a 100 mM stock in isopropanol.
- Stocks for adjusting chromatin samples to RIPA conditions: 10 % Triton X-100, 1 % SDS, 1 % sodium deoxycholate, 1.4 M NaCl.
- LiCl buffer: 250 mM LiCl, 10 mM Tris-HCl pH 8.0, 1 mM EDTA, 0.5 % NP-40, 0.5 % sodium deoxycholate (on ice).
- TE: 10 mM Tris-HCl pH 8.0, 1 mM EDTA.

Method

1. Thaw a 500 µl aliquot of chromatin and adjust to RIPA buffer conditions by sequential addition of 100 µl 10 % Triton X-100, 100 µl 10 % sodium deoxycholate, 100 µl 1 % SDS and 100 µl 1.4 M NaCl. Allow 2 min gentle mixing between additions for equilibration of the chromatin into the new conditions. Finally add 10 µl 100 mM PMSF. Always include one additional sample to be mock treated as a negative control.

2. Add 30 - 40 µl of the 50 % v/v PAS suspension to the chromatin sample. Incubate for 1 h at 4 °C, before removing the PAS by centrifugation in a microfuge at top speed for 30 sec. This acts as a preclearing step to reduce nonspecific binding to the protein A Sepharose.

3. Remove the chromatin sample to a new tube, and add 2 - 5 µg antibody. Incubate overnight at 4 °C, with gentle mixing. The optimal amount of antibody may need to be determined empirically, and a control (mock) IP without antibody should be carried out in parallel. Mock IP isolates DNA non-specifically, but specific antibody-bound DNA fragments should be several-fold enriched in the antibody IP.

4. Purify immunocomplexes by adding 30 - 40 µl 50 % v/v PAS, and incubating for 3 h at 4 °C, with gentle mixing.

5. Wash PAS-antibody-chromatin complexes 5 x for 10 min each in RIPA buffer, once in LiCl buffer, and twice in TE. All wash steps should be carried out at 4 °C using 1 ml wash buffer, and between washes centrifuge at full speed for 20 sec. to pellet the PAS before removing the supernatant.

6. Resuspend PAS complexes in 100 µl TE, add RNase A to 50 µg/ml and incubate for 30 min at 37 °C.

7. Adjust samples to 0.5 % SDS, 0.5 mg/ml proteinase K and incubate overnight at 37 °C, followed by 6 h at 65 °C.

8. Phenol-chloroform extract the sample, and back-extract the lower phenol phase by adding an equal volume of 50 mM Tris-HCl pH 8.0, mixing and centrifuging. Combine the aqueous phases from the phenol extraction the back-extraction, and chloroform extract.

9. Precipitate by adding 1 µl 20 mg/ml glycogen (as carrier), 1 /10 volume 3 M sodium acetate pH 5.2 and 2 volumes ice-cold 100 % ethanol. Store on ice for 30 min, before centrifuging at

4 °C for a further 30 min. Wash the DNA pellet in 70 % ethanol, air dry and resuspend in 20 µl dH_2O water. Store at -20 °C.

To minimise the risk of contamination during the PCR, we would recommend the use of aerosol-free pipette tips, and the storage of nucleotides, linkers, primers etc. in small aliquots to prevent contamination of valuable reagents (in addition to preventing frequent freeze-thawing which may cause destabilisation/inactivation of buffers).

Protocol 4. PCR amplification of immunoprecipitated DNA

Equipment and Reagents

- 5 mg/ml RNase A (DNase-free).
- 10 % SDS.
- 10 mg/ml proteinase K (in TE, stored at -20 °C).
- 1 µM linker DNA: Two oligonucleotides annealed: (i) a 24-mer of sequence 5´-AGA AGC TTG AAT TCG AGC AGT CAG (phosphorylated at 5´-end); (ii) a 20-mer of sequence 5´-CTG CTC GAA TTC AAG CTT CT. Store in small aliquots at -20 °C.

To produce the linker: 1) phosphorylate the 24-mer primer as follows.
1 µl of 1 mM 24-mer oligonucleotide
16 µl H2O (to 20 microL final volume)
2 µl of 10X T4 kinase buffer (Boehringer)
1 µl of T4 polynucleotide kinase 10 U/microL
incubate 1h at 37°C
incubate 10 min at 68°C (heat inactivate)
add 40 µl H2O, mix. add 240 µl 5M Ammonium acetate, mix. add 750 µl ethanol 80%, precipitate at -20°C, centrifuge 30 min, remove supernatant, dry on air and resuspend in 100 µl TE to obtain a 10 µM solution of Phosphoprimer.
2) generate the linker:
100 µl of 10 microM Phosphoprimer
100 µl of 10 microM primer
incubate 5min at 70°C (remove 2ndary structures, I used a thermoblock with agitation)
incubate 5min at 55°C (annealing)
let cool down slowly at room temperature to obtain 200 microL of 5 microM blunt PCR linker. This can be stored at -20°C and make 1 microM aliquots of 25 microL, by diluting to 1 microM with TE.

- 10 x ligation buffer: 0.5 M Tris-HCl pH 7.6, 125 mM $MgCl_2$, 250 mM DTT, 12.5 mM ATP Store at -20 °C in small aliquots.
- T4 DNA ligase, 4 U/µl.
- Taq polymerase and buffer (Boehringer).
- 10 mM dGTP, dATP, dCTP, dTTP (diluted from commercially available stocks in 10 mM Tris-HCl pH 8.0, and stored in small aliquots at -20 °C).

- PCR primer: linker oligonucleotide (ii), at 100 µM, stored in small aliquots at -20 °C.
- Phenol-chloroform, chloroform, 3 M NaOAc pH 5.2, 100 % ethanol (on ice), 20 mg/ml glycogen (Boehringer), 70 % ethanol, TE.
- HinD III restriction endonuclease and corresponding buffer.
- PCR purification columns (e.g. Qiagen).

Method

1. Ligate linker to 7 µl of the sample from *Protocol 3* by adding 1 µl 10 x ligation buffer, 1 µl 1 µM linker and 1 µl (4 U) T4 DNA ligase. Incubate overnight at 4 °C.

2. Amplification is carried out directly, without purification of the ligated DNA. Make the volume of the sample up to 78.5 µl with dH_2O, and add 10 µl 10 x Taq polymerase buffer, 2.5 µl each 10 mM dNTP, 1 µl 100 µM primer and 0.5 µl Taq polymerase. Use the following amplification scheme: 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; 1 cycle of 94 °C for 1 min, 55 °C for 1 min, 72 °C for 10 min.

3. Check 5 µl of the amplified samples on a 1% agarose gel; the product should be a smear ranging between 200 - 500 bp. Obvious bands in the smear may be a result of contamination.

4. Purify the amplified DNA by phenol-chloroform extraction and ethanol precipitation. Remove linker DNA sequences by digestion with 10 U HinD III. Purify the amplification products from linker DNA using Qiagen PCR purification columns, according to the manufacturer's conditions. The expected yield from the PCR reaction is approximately 5 µg, and amplification of DNA from the mock IP should be as efficient as that of the antibody IP.

5. Analysising the enrichment of putative target sequences in the PCR-amplified DNA.

5.1 Slot-blot analysis of the enrichment of putative PC target sequences.

When putative target sequences of the chromatin protein under study are known, it is possible to test whether they are enriched in the complex DNA sample obtained by ip from the cross-linked chromatin using slot-blot analysis. 100 ng of mock IP and antibody IP DNAs are immobilised on the membrane and hybridised with a radioactive probe from the DNA fragment to be tested (*Protocol 5A*). Relative enrichments are calculated as a ratio between antibody and mock IPs. In *Figure 2A*, relative enrichment is shown for a known Pc target site, a 6.0 kb EcoRI fragment from the bithorax complex (BX-C), named MCP (10, 11). A strong enrichment of this sequence in the Pc IP compared to the mock IP was seen in two independent *Drosophila* lines, as previously reported for *Drosophila* cultured SL-2 cells (4, 6). Only specific Pc target sites are enriched by this procedure (*Figure 2B*). Heat shock sequences, which are known not to be targets of Pc-binding, are not enriched. It should be noted that, although it is very valuable to have quantitative measurements of relative enrichments, it is not appropriate for scanning binding to extended chromosomal domains.

5.2 Mapping DNA target sites for Polycomb and GAGA factor in the *Drosophila* bithorax complex.

In order to scan binding to a large genomic region, or to compare relative binding affinity to neighbouring DNA fragments from regions with a size of several kb, the PCR-amplified DNA can be used as a complex radioactive probe to hybridise a membrane obtained by Southern blotting of an agarose gel containing a restriction digest of DNA from the corresponding genomic region. When large regions are analysed, genomic walks of several lambda phage or P1 clones are digested and loaded side by side on the gel. The restriction digests can be set to obtain DNA fragments as small as 500 bp, the maximal resolution of the methodology. *Figure 3* shows an analysis of Pc and GAGA factor binding to three regions from the BX-C. The first is the 6.0 kb MCP fragment mentioned previously. The second is another Pc target, a 3.4 kb EcoRI fragment containing the so-called BXD Polycomb response element (PRE) (12). The third sequence is a 7.6 kb fragment, named Peak C, which was identified during the analysis of the whole 340 kb BX-C in SL-2 cultured cells as a specific target for both PC and GAGA binding (6). Specific enrichments were observed for binding of both proteins to the same DNA subfragments in *Drosophila* embryos as previously described for SL-2 cells. These enrichment can be plotted in order to better display relative binding affinities to neighbouring regions (see *Figure 3D to F*). Importantly, the binding profile of the two proteins is qualitatively different. Pc, a chromatin-associated protein which associates to large regions of its target genes, shows rather broad peaks. On the other hand GAGA factor, a DNA binding protein which recognizes GA repeats on the DNA *in vitro* and *in vivo*, binds only to DNA fragments containing its target sites. Therefore, the methodology presented here is suited to mapping the binding of both broadly distributed proteins as well as sequence-specific binding proteins.

Protocol 5. Analysis of the enrichment of specific DNA sequences in PCR-amplified DNA.

Equipment and reagents

- Standard materials for agarose gel electrophoresis and Southern blotting onto positively charged nylon membrane (e.g. Genescreen Plus, NEN free call Nr France 0800/907762, Catalog-Nr: NEF976).
- Slot-blot apparatus minifoldII (Schleicher & Schuell, SRC 072/0)
- Phosphorimager apparatus.
- Denaturation buffer: 0.5 M NaOH, 1 M NaCl.
- Dilution buffer: 0.1X SSC solution, 0.125 M NaOH.
- Neutralisation buffer: 0.5 M Tris-HCl pH 7.5, 0.5 M NaCl.
- Standard materials for random prime DNA labelling with $\alpha(^{32}P)$-dATP (Specific activity 3000 Ci/mmol, Amersham).
- Hybridisation buffer: 0.5 M $NaHPO_4$ pH 7.2, 7 % SDS, 1 mM EDTA pH 8.0, 1 % BSA. A 1M $NaHPO_4$ pH 7.2 stock is 0.5 M Na_2HPO_4 containing 4 ml orthophosphoric acid per litre.

- Wash buffer 1: 40 mM $NaHPO_4$ pH 7.2, 5 % SDS, 1 mM EDTA pH 8.0, 0.5 % BSA.
- Wash buffer 2: 40 mM $NaHPO_4$ pH 7.2, 1 % SDS, 1 mM EDTA pH 8.0, 0.5 % BSA.

Method

A. Analysis by slot-blot of enrichment of specific DNA fragments

1. Denature 100 ng DNA of the mock IP and antibody IP samples by addition of 15 µl of denaturation buffer and incubation for 10 min at r.t.

2. Add 270 µl of dilution buffer and transfer on ice. Load samples in the slots of the minifold apparatus containing hybridisation membrane and let the liquid flow through the membrane for 30 min. Connecting the minifold to a vacuum pump until no liquid is left in the slot. Uncast the slot-blot and neutralise the membrane for 1-2 min in neutralisation buffer. Fix the DNA to the membrane by baking at 80°C for 1 h.

3. Prehybridise for 2 - 4 h at 65 °C in hybridisation buffer (hybridisation is carried essentially as described in ref 13.).

4. Label 25 ng of the specific DNA fragment to be tested for specific enrichment by random priming with 50 µCi α-(^{32}P)-dATP.

5. Denature the probe by boiling for 5 min, and cool in ice, before adding to 5- 10 ml hybridisation buffer and incubating overnight at 65 °C.

6. Wash the filter twice for 10 min at 65°C in wash solution A, and 3 - 4 x for 5 min each at 65°C in wash solution B. Expose the filter overnight to a Phosphorimager screen. Quantify the signals of the mock- and the antibody IP samples in order to calculate the specific enrichment due to IP of the cross-linked chromatin for the sequence used as probe .

7. After exposure, strip the probe from the membrane according to manufacturer's instructions, and reprobe with 25 ng of labelled *Drosophila* genomic DNA. This gives a precise estimate of the amount of DNA loaded on the membrane, to use as a correction factor for standardizing the calculated specific enrichments.

B. Searching target DNA binding sites by analysis of the Southern hybridisation profile to large genomic walks or subcloned DNA fragments

1. Prepare a Southern filter of a genomic walk which is a potential target of the immunoprecipitated protein, or with the subcloned putative target DNA fragments digested by appropriate restriction enzymes. Prehybridise for 2 - 4 h at 65°C in hybridisation buffer.

2. Random prime label 50 - 100 ng of amplified DNA with α-(^{32}P)-dATP.

3. Hybridise as described above (see A, steps 3 - 6), and expose the filter (overnight or several days) to a Phosphorimager screen.

4. Quantify the signal intensity of each band using the software package of the Phosphorimager, and calculate the enrichment of each fragment taking into account the following points:
(i) Intensity is proportional to M.W., and the resulting values should be normalised with respect to M.W. if the relative enrichments of different DNA fragments are being compared. The amount of signal per kb of DNA in each fragment may be calculated and plotted against a map of the genomic region (as shown for PC and GAGA factor in *Figure 3D to F*).
(ii) the mock IP probe should hybridise approximately uniformly to all fragments (dependent on M.W.). As the method is only semi-quantitative some sequence-specific differences in

amplification may occur, but amplification of different fragments generally varies by no more than 50 % from the mean. This degree of error must therefore be assumed for all experiments. If only a few random restriction fragments of a genomic walk hybridise to the mock IP probe, it is likely that too little input DNA was added to the ligation reaction, or that ligation occurred at low efficiency.
(iii) Repetitive elements are always strongly enriched in immunoprecipitations and therefore hybridise strongly to all ip DNA probes. These elements can be identified by their strong hybridisation to genomic DNA.
(iv) When analysing large genomic walks, antibody IP probes will hybridise to all fragments of a genomic region to some extent (although enriched fragments hybridise much more strongly). In order to accurately determine the background level, slot blot analysis may be used. Typically 100 - 200 ng of DNA from mock and antibody ips is immobilised on nylon membrane by slot blot and hybridised to a number of probes derived from the target DNA of interest. The resulting signals are quantitated and the actual enrichment accurately determined. Comparison between a number of fragments allows the setting of a "background" level, and only hybridisation signals above this level are considered to be enriched.

6. Concluding remarks

The formaldehyde cross-linking method presented here allows *in vivo* mapping of the DNA target sites of chromatin-associated proteins and site-specific DNA binding proteins in *Drosophila* embryos. This has several applications in the molecular dissection of regulation of gene expression for developmentally relevant genes. For example, whenever the potential target sites of a given regulatory factor are known, it can be determined whether they are actually bound and the developmental timing of DNA binding can also be analysed. Moreover, competition between regulatory proteins within the nucleus can be analysed, either at natural target sites or in transgenic constructs. In several regulatory sequences of developmental genes DNA binding sites for regulatory transcription factors are overlapping. With cross-linking analysis, it might be possible to investigate whether binding of a given regulator physically displaces a second from the DNA.

Many chromatin-associated proteins show no site-specific DNA binding *in vitro*. In this case, formaldehyde cross-linking appears to be the method of choice for analysis of binding to regulatory regions of potential target genes. Moreover, this method is also applicable to the analysis of histone modifications in the chromatin template *in vivo*. Formaldehyde cross-linking has already been successfully applied in the study of changes in histone acetylation in the mating type loci and telomeres of yeast chromatin (14). Additionally other chromosomal processes, e.g. DNA replication, recombination and repair, can now be subjected to an *in vivo* analysis of participating components.

References

1. McGhee, J.D. and P.H. von Hippel, Biochem., 1975. **14**: p. 1281-1296.

2. McGhee, J.D. and P.H. von Hippel, Biochem., 1975. **14**: p. 1297-1303.

3. Solomon, M.J. and A. Varshavsky, Proc. Natl. Acad. Sci. USA, 1985. **82**: p. 6470-6474.

4. Orlando, V. and R. Paro, Cell, 1993. **75**: p. 1187-1198.

5. Hecht, A., S. Strahl-Bolsinger, and M. Grunstein, Nature, 1996. **383**: p. 92-06.

6. Strutt, H., G. Cavalli, and R. Paro, EMBO J, 1997. **16**: p. 3621-3632.

7. Paro, R., Trends Genet., 1995. **11**: p. 295-297.

8. Farkas, G., *et al.*, Nature, 1994. **371**: p. 806-808.

9. Orlando, V., H. Strutt, and R. Paro, METHODS: A Companion to Methods Enzymol., 1997. **11**: p. 205-214.

10. Busturia, A. and M. Bienz, EMBO J., 1993. **12**: p. 1415-1425.

11. Zink, D. and R. Paro, EMBO J., 1995. **14**: p. 5660-5671.

12. Chan, C.S., L. Rastelli, and V. Pirrotta, EMBO J., 1994. **13**: p. 2553-2564.

13. Church, G.M. and W. Gilbert, Proc. Natl. Acad. Sci. USA, 1984. **81**: p. 1991-1995.

14. Braunstein, M., *et al.*, Genes Dev., 1993. **7**: p. 592-604.

15. Martin, C.H., *et al.*, Proc. Natl. Acad. Sci. USA, 1995. **92**: p. 8398-8402.

Figure Legends

Figure 1. Analysis of cross-linking efficiency with different concentrations of formaldehyde.
Cross-linking of embryonic chromatin was performed with 1.8% (A, C), or 1% (B, D) formaldehyde. (A) and (B): Analysis of the DNA from each fraction collected from a CsCl gradient of 1.8% and 1% formaldehyde cross-linked chromatin, respectively. On top of the gel, fraction numbers are indicated. On the right of each panel the size of the DNA M.W. markers (Boehringer) is shown in bp. In the 1.8% sample, the majority of the DNA elutes with a density of 1.36 to 1.42 g/cm^3, corresponding to protein-DNA complexes, while in the 1% sample, a large amount of the DNA elutes in the bottom of the gradient. This has a density of about 1.60 g/cm^3, corresponding to free, uncross-linked DNA.
(C) and (D): The gels from (A) and (B), respectively, were photographed, and the intensity of the DNA signal was estimated by scanning each lane using NIH Image software package. The relative intensity of the signal (after background subtraction, calculated from an empty lane) is plotted against the fraction number. The density of fractions which were pooled for chromatin analysis of the 1.8% sample is indicated.

Figure 2. The relative enrichment of the MCP sequence by Pc immunoprecipitation in wild type and transgenic fly lines.
Cross-linked chromatin from two *Drosophila* lines was analysed: wild type, OregonR flies (wt) and the 5F24 25,2 transgenic line, which contains a P element-derived transgene inserted in the X chromosome (11). The enrichment factor of Pc IP versus Mock IP chromatin, calculated by Phosphorimager quantitation of the hybridisation signals, is reported to the right of each panel.

(A) Slot-blot enrichment of the 6.0 kb EcoR I MCP element from the BX-C was assayed by hybridisation to 100 ng of cross-linked, PCR-amplified DNA obtained from mock or Pc IP. A substantial enrichment factor is observed in both lines.
(B) The same DNAs were hybridised to a 10 kb EcoR I DNA fragment from the Heat shock 87C genomic locus (4). In this case, no significant enrichment is detected.

Figure 3. Hybridisation of control, Pc, and GAGA immunoprecipitated DNA to putative target sites in the BX-C.
From the previously published analysis of Pc and GAGA factor binding to the BX-C, three fragments subcloned into Bluescript KS+ (6) were analysed in order to characterize binding of Pc and GAGA at higher resolution. The three fragments are (i) a 3384 bp EcoR I fragment, from coordinate 218 241 to 221 625 of the published BX-C sequence (15). This fragment contains a Pc and GAGA factor target site, named the BXD PRE (12). (ii) a 7652 bp EcoR I fragment (coordinates 123 772-131 424), a PC and GAGA target site named Peak C (6). (iii) a 5989 bp EcoR I fragment (coordinates 109 688-115 677) containing the MCP element, a known Pc target (6, 10).
(A-C). DNAs were digested, run on a 1% agarose gel, blotted and hybridised to PCR amplified mock IP (panels A, B, C, left), GAGA IP (panel A, right) or PC IP (panels B and C, right). On the right side of each panel, the size of the DNA fragments (in bp) most strongly bound by Pc and GAGA factor is shown. Panels A and B, lanes 1 and 4; BXD PRE digested by Kpn I (K) and Pst I (P). Panels A and B, lanes 2 and 5; Peak C digested by Pst I, BamH I (B) and Xho I (Xh). Panels A and B, lanes 3 and 6; MCP digested by Acc I (A) and Pst I. In order to analyse Pc binding to the two largest MCP fragments of 2208 bp and 1556 bp, the 2208 bp fragment was gel-eluted and digested with Rsa I (R) and Xmn I (X) (C, lanes 8, 10), while the gel-eluted 1556 bp fragment was digested by Pvu II (Pv) (C, lanes 7, 9).
(D-F). Calculated relative enrichments for each DNA fragment are shown on the right of each panel. For BXD PRE and Peak C (D, F, respectively), enrichments for Pc and GAGA factor are shown, whilst only Pc enrichments are shown for MCP (E) since no GAGA factor binding could be detected to any fragments of the MCP region (panel A, compare lane 3 with 6). The plots show the DNA fragments of each DNA element to scale in the proximal-distal orientation along the BX-C.

Figure 1

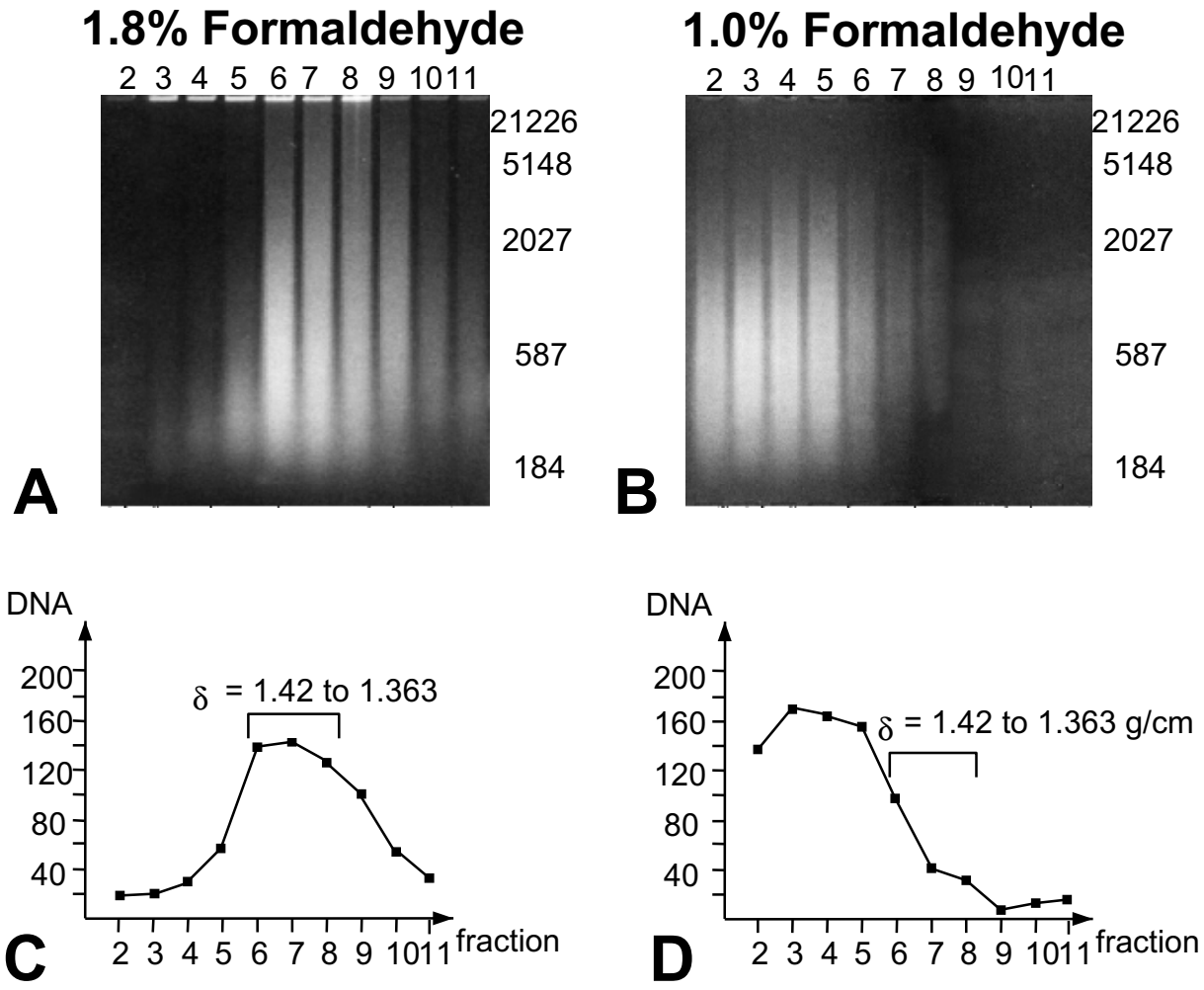

1.8% Formaldehyde
2 3 4 5 6 7 8 9 1011
21226
5148
2027
587
184
A
1.0% Formaldehyde
2 3 4 5 6 7 8 9 1011
21226
5148
2027
587
184
B
DNA
200
160
120
80
40
δ = 1.42 to 1.363
2 3 4 5 6 7 8 9 1011
fraction
C
DNA
200
160
120
80
40
δ = 1.42 to 1.363 g/cm
2 3 4 5 6 7 8 9 1011
fraction
D

Figure 2

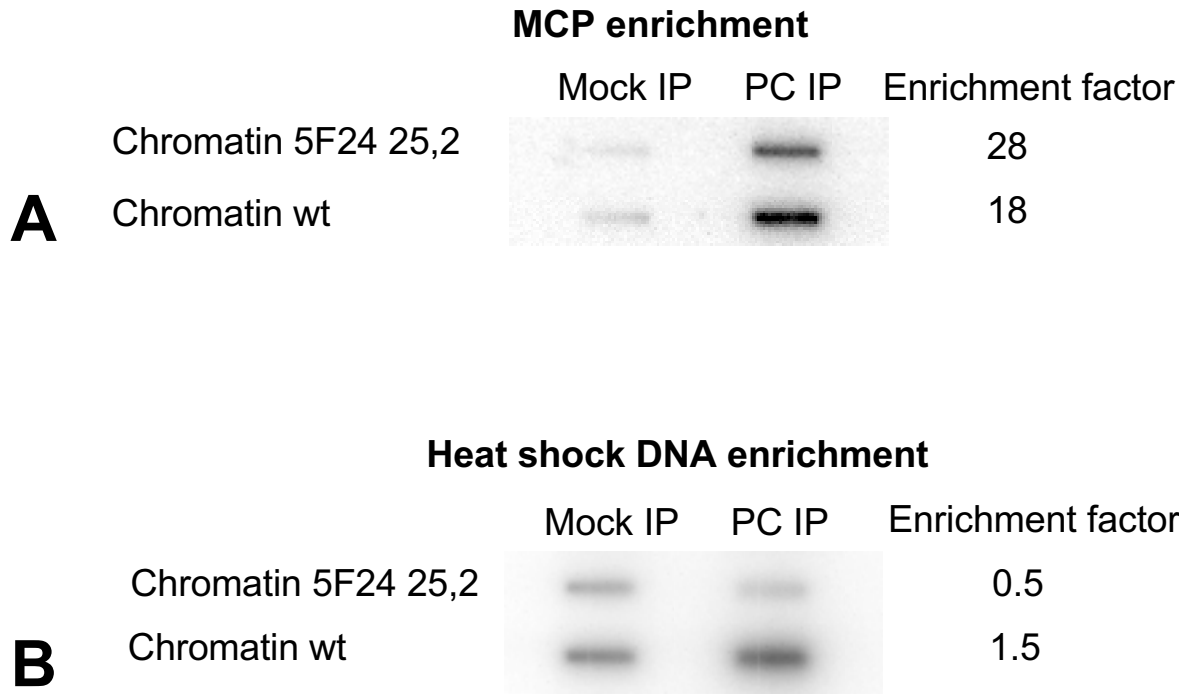
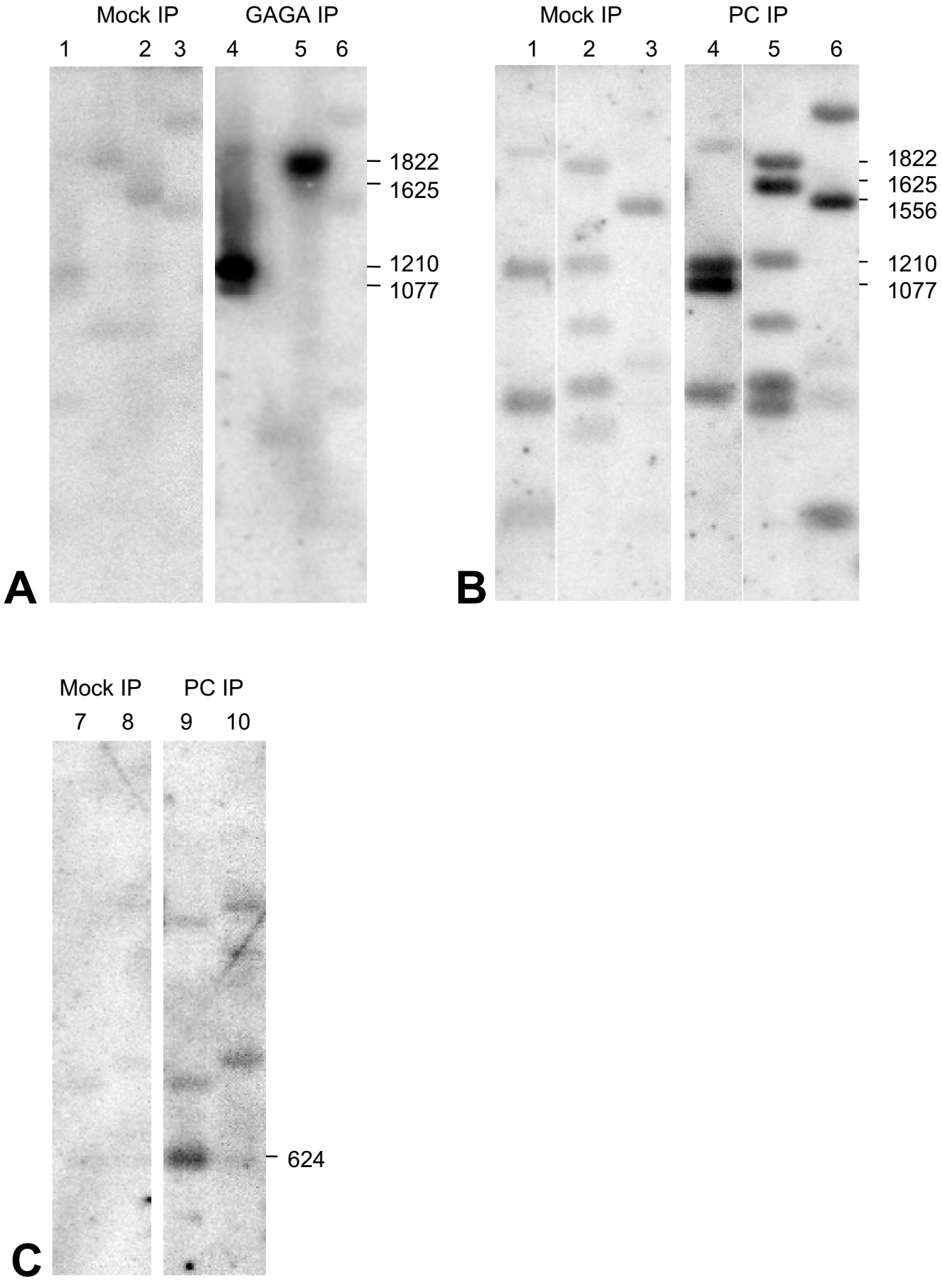
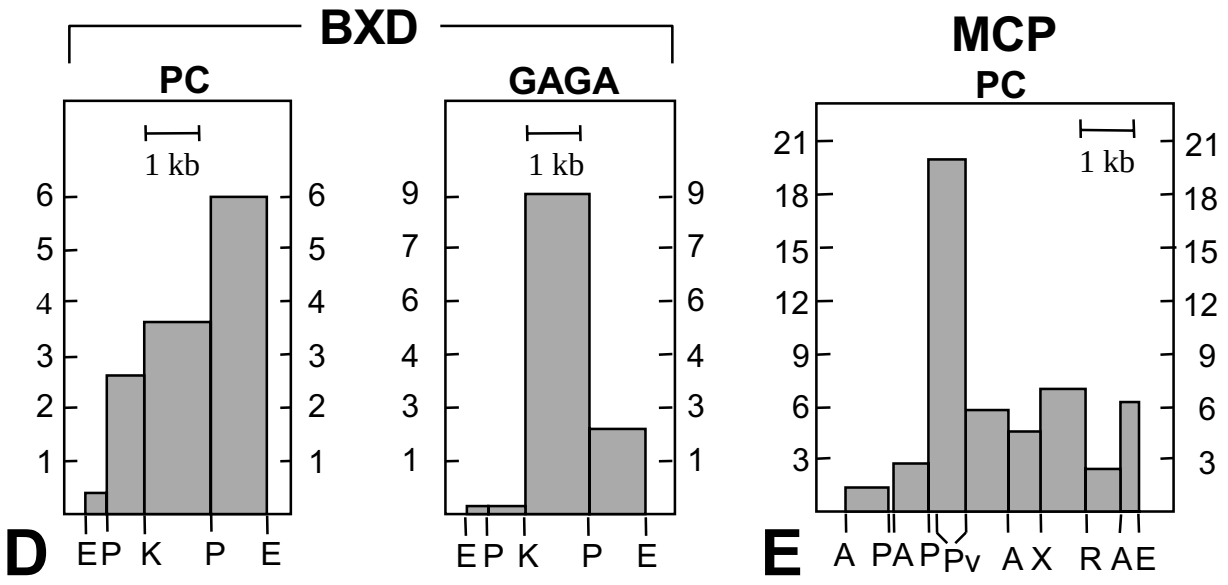
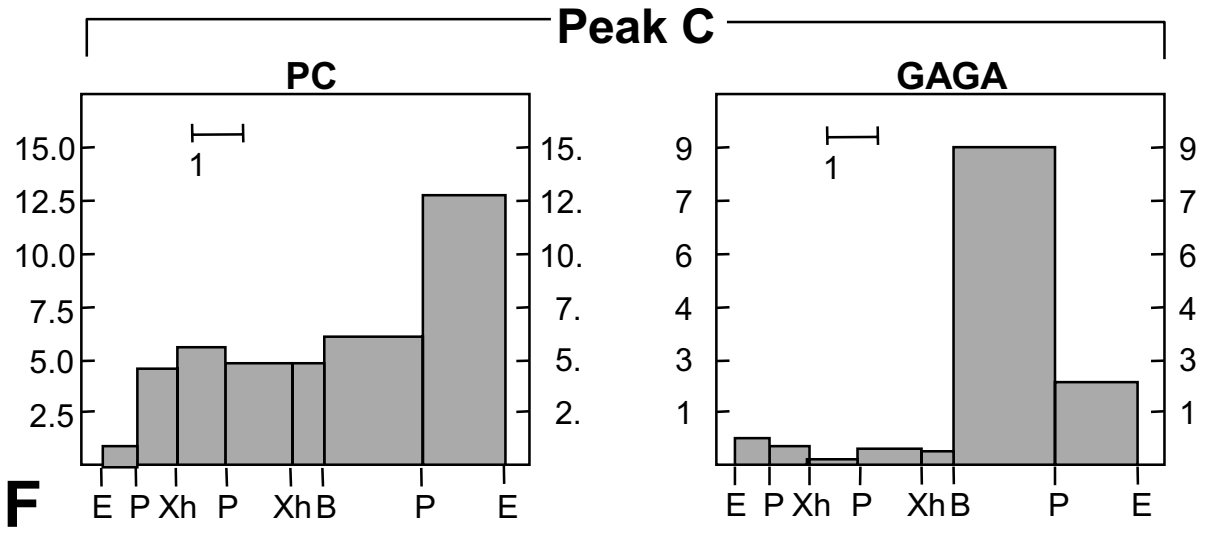

MCP enrichment
Mock IP
PC IP
Enrichment factor
Chromatin 5F24 25,2
28
A
Chromatin wt
18
Heat shock DNA enrichment
Mock IP
PC IP
Enrichment factor
Chromatin 5F24 25,2
0.5
B
Chromatin wt
1.5

Figure 3


Mock IP
GAGA IP
1
2
3
4
5
6
1822
1625
1210
1077
A
Mock IP
PC IP
1
2
3
4
5
6
1822
1625
1556
1210
1077
B
Mock IP
PC IP
7
8
9
10
624
C


BXD
PC
GAGA
1 kb
1 kb
6
5
4
3
2
1
9
7
6
4
3
1
E P K P E
E P K P E
D
MCP
PC
1 kb
21
18
15
12
9
6
3
A PA P Pv A X R A E
E


Peak C
PC
GAGA
1
1
15.0
12.5
10.0
7.5
5.0
2.5
9
7
6
4
3
1
E P Xh P Xh B P E
E P Xh P Xh B P E
F