METHODS: A Companion to Methods in Enzymology 11, 205–214 (1997)
Article No. ME960407

Analysis of Chromatin Structure by *in Vivo* Formaldehyde Cross-Linking

Valerio Orlando,[1] Helen Strutt, and Renato Paro[2]

ZMBH, Center for Molecular Biology, University of Heidelberg, Im Neuenheimer Feld 282, 69120 Heidelberg, Germany

Recent advances leave no doubt that higher order chromatin structures play a fundamental role in many developmentally important mechanisms of gene regulation. In particular analyses in genetic model systems like yeast and *Drosophila* uncovered novel proteins that are involved in the regulation of chromatin structures. Many of these proteins do not bind directly to DNA but interact in large multimeric complexes. To identify the DNA elements regulated by these multiprotein complexes, alternative approaches to the standard methods of DNA–protein analysis had to be devised. Here we present a method that preserves the architecture of the higher order chromatin structures by cross-linking cells *in vivo* with formaldehyde. An immunoprecipitation strategy is then used to identify the DNA targets of chromosomal proteins of interest. This method can be applied to study the distribution of proteins at high resolution over extended chromosomal regions. © 1997 Academic Press

Formaldehyde is a powerful, high-resolution (2 Å), and easily reversible cross-linking agent that efficiently produces both protein–DNA, protein–RNA, and protein–protein cross-links *in vivo* (1–8). Thus protein–DNA interactions in virtually intact chromatin can be analyzed by this method. Within minutes cross-linking results in the formation of biopolymers, thus preventing any large-scale redistribution of cellular components. Formaldehyde does not react with free double-stranded DNA, avoiding kinetic constraints due to DNA damage. In addition, formaldehyde cross-links can be reversed under mild conditions so that DNA and proteins can be further analyzed, separately.

Figure 1 shows an outline of the method. Living cells are fixed with formaldehyde. After sonication, fixed chromatin is purified by cesium chloride gradient centrifugation. The chromatin fragments containing the protein of interest are purified by selective immunoprecipitation. Finally, cross-links are reversed and the specifically enriched DNA is analyzed. The experimental scheme originates from the pioneering work of Alexander Varshavsky (MIT, Cambridge) who developed the chromatin fixation, purification, and immunoprecipitation scheme for analyzing the distribution of histones in *Drosophila* heat-shock gene promoters (7). We have substantially broadened the potential of the method by adapting it to the analysis of general transcription factors and of protein components of multimeric complexes associated with the chromatin (Ref. 8; Orlando, V., Muchardt, C., Yaniv, M., and Paro, R., in preparation). We have devised a PCR and hybridization strategy that allows the identification of all binding sites of a particular protein within a genomic region of interest. In particular, the DNA associated with specifically immunopurified chromatin is amplified by PCR and used as the probe in a Southern analysis of a given genomic region. This type of direct mapping can lead to the identification of binding sites over extended genomic regions with a resolution of 1 kb. The possibility of identifying *in vivo* target sequences in chromatin can be of advantage for diverse applications, including high-resolution mapping and target gene identification (9).

DESCRIPTION OF THE METHOD

1. *In Vivo* Formaldehyde Fixation of Cells

Approximately 1×10^9–5×10^9 tissue culture cells from the *Drosophila melanogaster* embryonic cell line SL-2 are grown at 25°C in spinner flasks in Schneider's *Drosophila* medium (Serva or GIBCO) supplemented with heat-inactivated 10% FCS, 100 U/ml penicillin, and 100 μg/ml streptomycin (GIBCO). To detect the chromatin-containing fractions in the cesium chloride gradient (see Section 4), the DNA is labeled with 1 μCi/ml [methyl-^{3}H]thymidine (specific activity 83 Ci/mmol, ICN), added 36–48 h before the cells are fixed.[3]

[1] Present address: DIBIT Scientific Institute, H. San Raffaele, Via Olgettina 58, Milano, Italy.

[2] To whom correspondence should be addressed. Fax: 49-6221-54 58 93.

[3] When using starting material other than tissue culture cells (e.g., embryos) the DNA in the gradient can be traced by reversing the cross-links (see Section 6) of a small aliquot of gradient fractions

1046-2023/97 $25.00
Copyright © 1997 by Academic Press
All rights of reproduction in any form reserved.

The chemistry of formaldehyde cross-links are described in Fig. 2. Living cells are fixed by adding directly to the growth medium one-tenth volume of 11% formaldehyde (HCHO) solution (prepared from a 37% HCHO/10% methanol stock solution; Merck), in 0.1 M NaCl, 1 mM Na-EDTA, 0.5 mM Na-EGTA, 50 mM Hepes, pH 8.0. Cells are transferred to 4°C and allowed to fix for up to 1 h. Preliminary time-course experiments spanning between 5 min and 1 hr of fixation should be performed to yield the best combination of *in vivo* fixed chromatin, high DNA recovery, and small average size of chromatin fragments. Cross-linking for 1 h was found to be the most suitable time for *Drosophila* SL-2 chromatin. An excess of cross-linking results in a substantial loss of material and a higher average size of DNA fragments after sonication (see Section 2). Figure 3 shows a comparison between an efficient cross-linking (a) and an inefficient cross-linking (b). When not sufficiently cross-linked, virtually all of the input DNA is found at the bottom of the gradient, whereas the cross-linked chromatin has a density of about 1.39 g/cm^3. For specific purposes cross-linking time can be considerably reduced or prolonged (a few minutes up to several days; Ref. 7). In our hands, cross-linking times longer than 6 h resulted in considerable (70%) loss of material and a high average size (20 kb) of the chromatin DNA fragments.

To stop the cross-linking reaction glycine is added to 0.125 M. Fixed cells are pelleted by centrifugation at 780*g* for 10 min at 4°C. The cells are rinsed with ice-cold PBS and washed sequentially for 10 min each in 15 ml of 0.25% Triton X-100, 10 mM Na-EDTA, 0.5 mM Na-EGTA, 10 mM Tris–HCl, pH 8.0, and in 15 ml of 0.2 M NaCl, 1 mM Na-EDTA, 0.5 mM Na-EGTA, 10 mM Tris–HCl, pH 8.0. The final pellets are resuspended in 8–16 ml of 1 mM Na-EDTA, 0.5 mM Na-EGTA, 10 mM Tris–HCl, pH 8.0, divided into 5-ml aliquots for sonication, and stored on ice. Degradative activities are not prominent in fixed material. Therefore the above-described washing steps can also be performed at room temperature.

2. Chromatin Solubilization by Sonication

Fixed cells are highly resistant to restriction enzyme digestion or DNase I treatment. Therefore, soluble chromatin can only be produced efficiently by mechanical shearing. As mentioned above, the extent of cross-linking is critical for the efficient disruption of the fixed cells and also affects DNA yield and the average size of chromatin fragments. It should be stressed that material that is over-cross-linked will not produce small chromatin fragments, even by prolonged sonication. Figure 4 shows two typical results obtained with Schneider cells cross-linked for 1 h. In lane 1, the DNA size ranges from 20 kb to 100 bp, whereas an optimal result is shown in lane 2, where the chromatin DNA is homogeneously distributed around 600 bp. Each cell type may require fine tuning of cross-linking time to prevent over-cross-linking. In general, the size of the DNA fragments may be critical for high-resolution mapping studies of binding sites. If the aim of the experiment is either to show binding of a protein to a particular site (for example upon appropriate induction) by slot blot analysis or to search for target genes (screening of genomic libraries), optimizing this parameter is not so important.

Prior to the final processing of the sample by a sonicator, variables such as processing time, output control setting, and tip immersion depth are estimated. In particular, the tip immersion depth should be adjusted, keeping in mind the following principles: (i) energy radiates only from the horn's tip and is most concentrated within 13 mm of the face of the tip; (ii) the energy pattern from the tip of the horn has a tendency to push away solid particles (in this case the cross-linked cells): if the tip is placed too deep in the liquid, most of the solid will not receive treatment, but simply spin or circulate around the container; (iii) violent motion on the surface and foaming result in loss of energy coupled to the solution. The addition of microglass beads (0.1–0.5 mm diameter) improves the shearing efficiency. A ratio of 1 part glass beads to 3 parts liquid is recommended.

We routinely use the Branson Model 250 sonifier with microtip at constant power. Each 5-ml sample is kept on ice–salt or ice–ethanol in a glass beaker put on a lab jack and the microtip appropriately immersed. The output control is slowly advanced to setting 5 (an output of 60–70 W) and the sample sonicated for 30 s. This treatment is repeated twice with 1-min intervals. Mechanical shearing is effective during these steps; however, longer treatments do not improve the disruption. In SL2 cells fixed for 1 h, this produces DNA fragments of an average size of less than 1 kb. The efficiency of sonication can be checked by examining 10 μl of the sample under phase-contrast microscopy. The sample should be devoid of intact cells or large particles.

3. Chromatin Purification by Isopycnic Centrifugation

Samples are adjusted to 0.5% Sarkosyl (sodium lauryl sarcosine) and gently swirled for 10 min at room temperature. Debris is eliminated by centrifugation for 5 min at 15,000*g*. Nonchromatin-bound cross-linked proteins, together with naked DNA and RNA, are eliminated by cesium chloride isopycnic centrifugation. Samples are adjusted to 1.42 g/cm^3 CsCl (567.8 mg/ml), brought to 5 ml with the same TE-sarkosyl buffer, and centrifuged in a Beckman SW55Ti rotor at 40,000 rpm

around the expected density values and running on an agarose gel. Locating the DNA-containing fractions by measuring the O.D. is not reliable. Cross-linked material has variable O.D. spectra and in addition each gradient fraction contains proteins and nucleic acids that could make the identification of the chromatin-containing fractions problematic.

for 72 h at 20°C. A refractometer may be used to check the exact density. It is important not to overload gradients; at the end of the run a broad sarkosyl/lipid/protein aggregate is visible at the top of the gradient. Virtually all the sarkosyl is eliminated in this way (10, 11). The larger the aggregate the higher is the chance that the chromatin will become trapped in it and eventually be lost. Normally, for the above-mentioned starting material, four gradients are run in parallel (approximately 200 μg DNA per gradient).

4. Collection and Identification of Chromatin Fractions

Fractions of 0.4 ml are collected from the bottom of the gradient using a 0.25-mm capillary connected to a

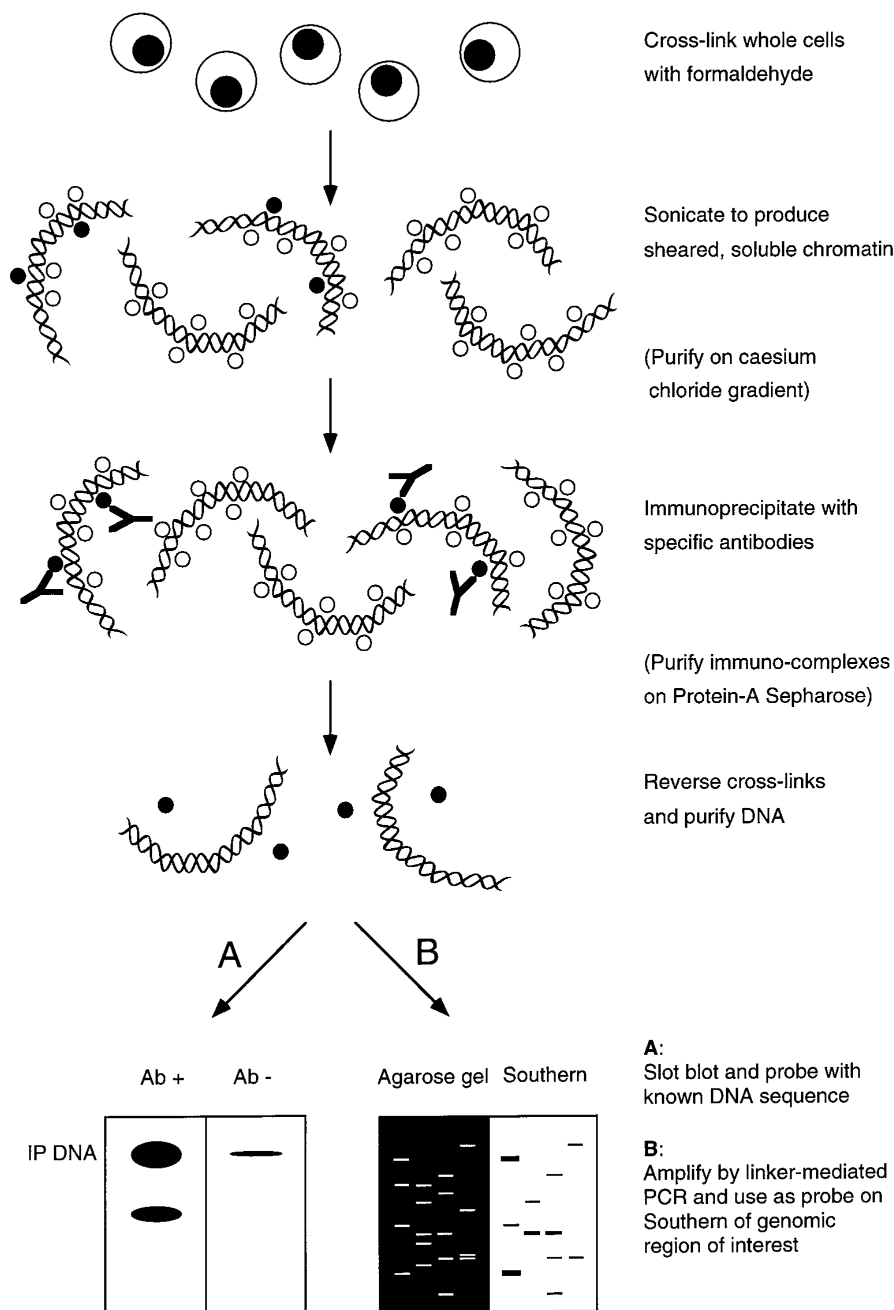

FIG. 1. Summary of the formaldehyde cross-linking and immunoprecipitation method. Whole cells are cross-linked with formaldehyde and soluble, sheared chromatin is isolated. Antibodies against a protein of interest are used to immunoprecipitate from chromatin and specifically bound DNA is purified. (A) Slot blot analysis. Immunoprecipitated DNA (IP DNA) from plus (Ab+) and minus (Ab−) antibody immunoprecipitations is immobilized on slot blot and probed with known genomic DNA sequences. (B) Southern analysis. After PCR amplification, immunoprecipitated DNA is used as a probe on a Southern blot.

peristaltic pump. A lab jack under the pump is convenient at this step to lower the capillary into the bottom of the centrifuge tube without turbulence. Ten microliters of each fraction is analyzed by scintillation counting, to identify peak DNA–chromatin-containing fractions (normally corresponding to ∂ = 1.39 g/cm^3; see Fig. 3). Three to four fractions are pooled and dialyzed overnight at 4°C against 5% glycerol, 1 mM Na-EDTA, 0.5 mM Na-EGTA, 10 mM Tris–HCl, pH 8.0, using micro-collodion bags (Sartorius). Fixed chromatin can be

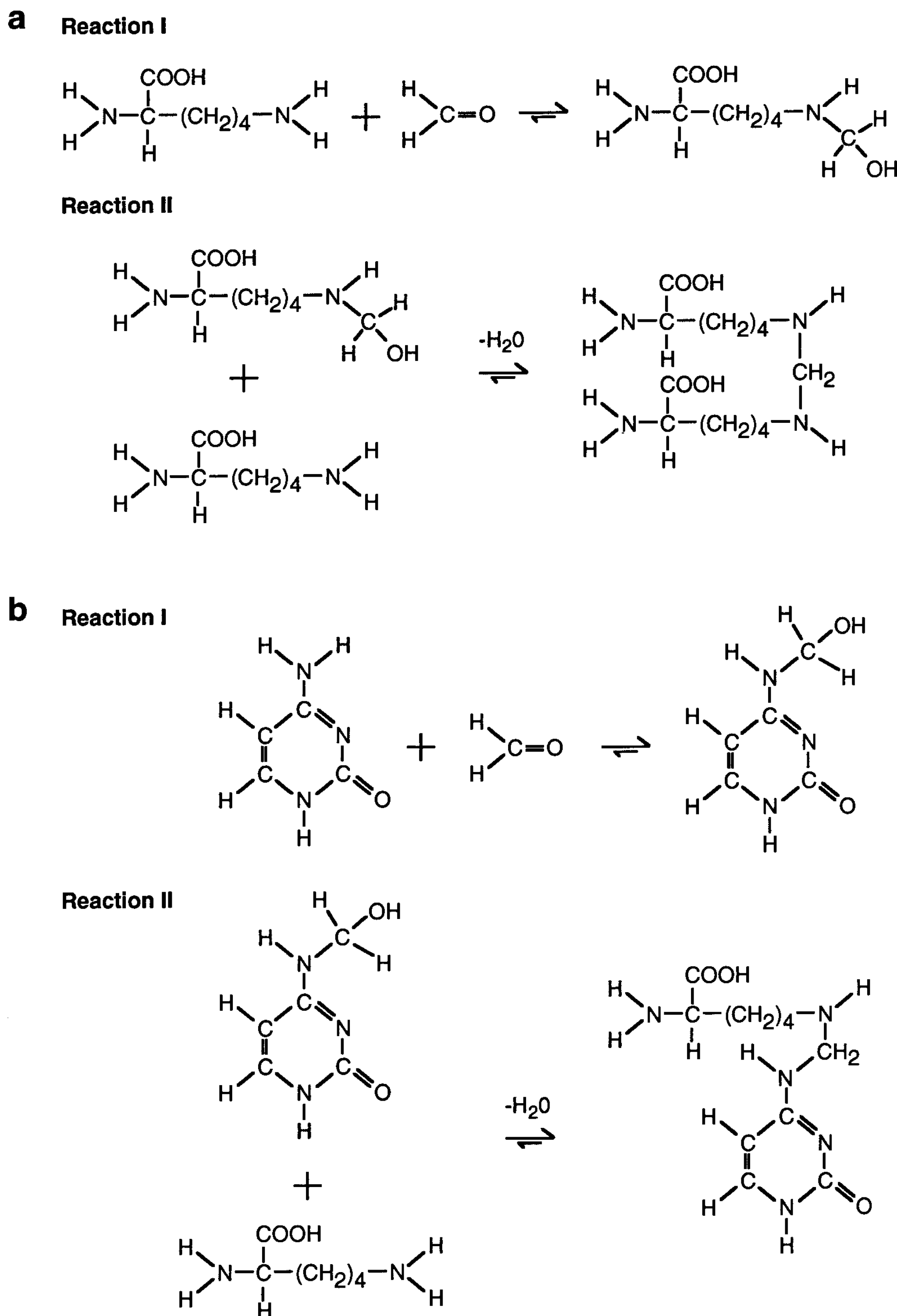

FIG. 2. Chemical cross-linking of DNA and proteins by formaldehyde. Formaldehyde (HCHO) is a very reactive dipolar compound in which the carbon atom is the nucleophilic center. Amino and imino groups of proteins (e.g., the side chains of lysine and arginine) and of nucleic acids (e.g., cytosine) react with formaldehyde, leading to the formation of a Schiff base (reaction I). This intermediate can react with a second amino group (reaction II) and condenses (19, 20). Cross-links may be reversed by heating in Tris–HCl-containing buffers. This leads to a drop in pH and protonation of amino groups, thus forcing the equilibrium in the reverse direction. (a) Formaldehyde-mediated cross-linking between the side chains of two lysines. (b) Cross-linking between cytosine and lysine.

stored for several months at −80°C in small aliquots suitable for immunoprecipitation. For our purposes aliquots of approximately 0.5 ml, containing 30–60 μg of DNA, were used. DNA content can be estimated by reversing the cross-links of a small aliquot of the purified chromatin (Section 6; Ref. 6).

5. Immunopurification of Cross-linked Chromatin

Prior to immunoprecipitation of fixed chromatin, it is absolutely necessary to test the antibody to check its compatibility with the detergents utilized. The conditions described below work well for rabbit polyclonal antibodies and are considered to be very stringent (7, 8). Interestingly, some proteins seem to be better or even only immunoprecipitated under high-detergent conditions. This is probably due to solubility problems circumvented in stringent buffers. Samples of fixed chromatin are adjusted to RIPA buffer (1% Triton X-100, 0.1% sodium deoxycholate, 0.1% SDS, 140 mM NaCl, 1 mM PMSF) by adding each component sequentially and mixing gently for 2 min between additions. As a preclearing step 20–60 μl of 50% (v/v) protein A–Sepharose CL4B (Pharmacia; equilibrated at 100 mg/ml in RIPA buffer) can be added, incubated for 1 h at 4°C, and removed by centrifugation at 12,000*g* for 1 min. A suitable amount of antibody (0.2–1 μg) is added to a fixed chromatin aliquot and incubated on a rocking platform overnight at 4°C. Excess antibody results in higher overall DNA yields, but a lower specificity of antibody binding, and thus lowers the relative enrichment of specifically immunoprecipitated DNA. In parallel, a control immunoprecipitation with no antibody should be carried out. Immuno-complexes are recovered by adding 10–50 μl of 50% (v/v) protein A–Sepharose to the sample and incubating with rocking for 3 h at 4°C. However, Protein A is not efficient in binding to mouse monoclonals (12). When mouse monoclonal antibodies are used, precoated beads should be avoided (e.g., rabbit anti-mouse IgGs) but rather specific magnetic dynabeads (Dynal) or GammaBind Plus (Pharmacia) should be used. Note that IgG-coupled Dyna-

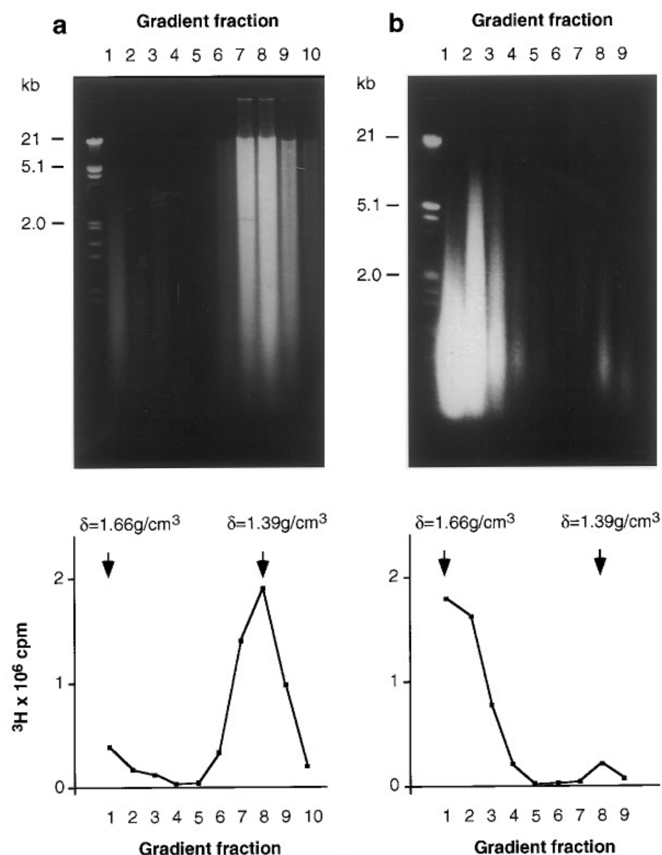

FIG. 3. Gradient profile of cross-linked DNA. Panel (a) is from cells that were well cross-linked panel (b) is from cells that were insufficiently cross-linked. After formaldehyde cross-linking, [methyl-^{3}H]thymidine-labeled cells were disrupted by sonication, and chromatin was purified on a cesium chloride gradient. Nine to ten gradient fractions were collected, and the total ^{3}H counts in each fraction were measured (bottom panels). Fraction 1 is from the bottom of the gradient and fraction 9 from the top. In (a) the majority of the DNA is in fractions 7–9 and is maximum in the fraction corresponding to a density of 1.39 g/cm^3, the position of DNA–protein complexes. In (b) most of the DNA is at the bottom of the gradient (density 1.66 g/cm^3), the position of noncovalently cross-linked DNA. The cross-links were reversed, and the DNA was purified and run on a 0.5% agarose gel (top panels). On the left side of each gel is a molecular weight marker (lambda *Eco*RI–*Hin*dIII), with various fragment sizes shown. In (a) the cross-linked DNA in fractions 7–9 has an average size of 2–3 kb, thus indicating that it is slightly over-cross-linked. In (b) the average size of DNA in fractions 1–3 is less than 500 bp, another indication of insufficient cross-linking.

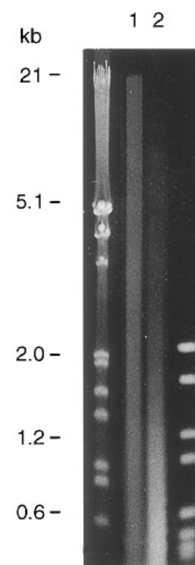

FIG. 4. The average size of chromatin fragments is dependent on the extent of cross-linking. To the left and right sides of the gel are molecular weight markers (lambda *Eco*RI–*Hin*dIII and pBR328 *Bgl*I/pBR328 *Hin*fI, respectively), and approximate fragment sizes are indicated. Lane 1 shows chromatin that is too large and lane 2 shows chromatin of an optimal average size (600 bp).

beads have a 10 times lower capacity than Sepharose-coupled ones.

Immunocomplexes are pelleted at 12,000*g* for 20 s, and the pellets are washed five times for 10 min each in 1 ml of RIPA buffer, once in 0.25 M LiCl, 0.5% NP-40, 0.5% sodium deoxycholate, 1 mM Na-EDTA, 10 mM Tris–HCl, pH 8.0, and twice in 1 mM Na-EDTA, 10 mM Tris–HCl, pH 8.0. Pellets are finally resuspended in 100 μl of 1 mM Na-EDTA, 10 mM Tris–HCl, pH 8.0.

6. Reversal of Cross-links and DNA Purification

DNase-free RNase at 50 μg/ml is added to the immunoprecipitated chromatin and incubated for 30 min at 37°C. To purify the immunoprecipitated DNA, samples are adjusted to 0.5% SDS, 500 μg/ml Proteinase K (Boehringer) and incubated for several hours (best overnight) at 37°C. A further treatment for 6 h at 65°C is necessary to complete the reversal of cross-links (6). Samples are extracted once with phenol–chloroform and the organic phase is back-extracted with 1 vol of 50 mM Tris–HCl, pH 8.0. Samples are then extracted once with chloroform–isoamyl alcohol; the DNA is precipitated in the presence of 20 μg glycogen as carrier and then resuspended in a suitable volume of buffer according to successive manipulations. Typically, from a 500-μl aliquot of chromatin containing 30–60 μg DNA, our immunoprecipitation experiments gave a yield of 1–10 ng purified DNA. For convenience, a chromatographic detection system (e.g., DNA DipStick Kit, Invitrogen) could be used for quantifying such small amounts of DNA. It is our experience that if too-large amounts of DNA are obtained from the immunoprecipitation (e.g., 1 μg out of 50 μg), nonspecific enrichments are expected.

In addition, the enrichment of a particular protein from cross-linked chromatin can be tested. In this case cells must be labeled with a pulse of radioactive amino acids before fixation and chromatin purification. After immunoprecipitation cross-links are reversed by boiling in SDS-denaturing buffer for at least 30 min. Samples are run directly on a denaturing gel and analyzed by autoradiography.

7. Analysis of Immunoprecipitated DNA and Identification of Binding Sites

There are two potential steps in the analysis of immunoprecipitated DNA. The first requires at least a known or suspected putative target sequence. In this case the specificity and efficiency of the immunoprecipitation can be tested by hybridizing this particular sequence to the bulk of the immunoprecipitated DNA in a slot blot experiment (7, 8, 13). Typically, half of the immunoprecipitated DNA would be immobilized on a nylon membrane by slot blot. Relative enrichment for a potential target sequence can be determined by comparing the hybridization of a specific probe to immunoprecipitated DNA, DNA from a control immunoprecipitation (with no antibody), or a known amount of genomic DNA. Figure 5 shows the results of such an analysis, in which the Mcp element, a target sequence for the Polycomb protein in the bithorax complex (BX-C) of *Drosophila,* is specifically enriched in Polycomb immunoprecipitations.

If no sequence is known or if the distribution of a regulatory factor over a large genomic region is the aim of the investigation, then the immunoprecipitated DNA itself can be used as a probe in a Southern analysis as described in Section 9. The latter case normally requires a PCR step that is described in Section 8. A PCR amplification step for slot blot analysis may also be necessary in the case of low immunoprecipitation efficiency (13).

8. Amplification of Immunoprecipitated DNA by Linker-Modified DNA PCR

To have enough DNA to be labeled and used as a probe in Southern analysis a linker-modified DNA PCR strategy is carried out. Immunoprecipitated DNA is digested with a restriction enzyme and ligated to an

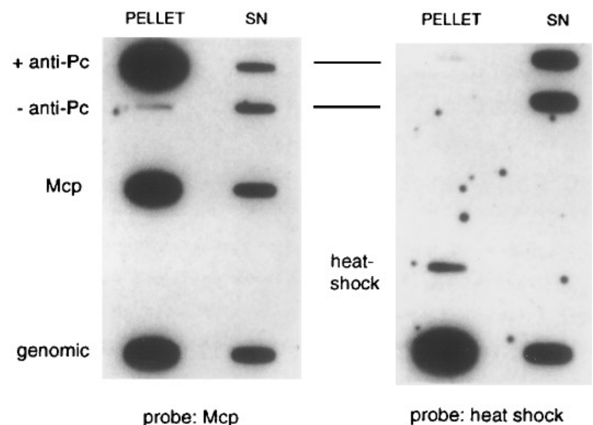

FIG. 5. Mcp sequences are enriched in immunoprecipitated Pc-containing chromatin. Slot blot analysis of the DNA obtained from immunoprecipitation of crosslinked chromatin with (+anti-Pc) and without (−anti-Pc) anti-Pc antibodies. (a) Equal amounts (~1 ng) of decross-linked purified DNAs from immunoprecipitations with and without anti-Pc antibodies (PELLET) were applied onto individual slots on a nylon filter in a slot blot apparatus and hybridized with the 5.2-kb *Eco*RI fragment covering the Mcp element of the bithorax complex of *Drosophila.* An aliquot (~100 ng) of the supernatants (SN) from the same immunoprecipitations, loaded aside as a control, shows no enrichment. As a positive control genomic equivalents of the 5.2-kb *Eco*RI Mcp fragment (70 and 7 pg) and 200 and 20 ng of SL-2 genomic DNA were loaded. (b) The same filter as described in (a) stripped and rehybridized with a 10-kb *Eco*RI fragment derived from the heat shock 87°C locus of *Drosophila.* No major enrichment for the heat shock sequences is observed. Due to an underloading of the 10-kb heat shock fragment, the signal in the positive control is weaker than expected. The weak signal with the heat shock probe in the +anti-Pc slot was due to incomplete stripping of the Mcp probe from the filter before rehybridization. (From V. Orlando and R. Paro, *Cell* **75,** 1187–1198, 1993, reprinted with permission of Cell Press.)

appropriate linker with a compatible end. Oligonucleotide primers homologous to the linker sequences are subsequently used to amplify the internal genomic sequences. It should be noted, however, that restriction fragments of different sizes may not amplify with equal efficiency. This effect must be taken into account if the relative distribution of a protein within a given genomic region is being investigated. If the genomic sequence is available, the distribution of restriction sites can be analyzed, particularly in terms of identifying very large fragments that may amplify poorly. To obtain accurate binding profiles, it may be necessary to compare the results obtained with immunoprecipitated DNA cut with a variety of restriction enzymes.

To prepare the adapter, two oligonucleotides, a 24-mer of sequence 5′GATCAGAAGCTTGAATTCGAGCAG and a 20-mer of sequence 5′CTGCTCGAATTCAAGCTTCT (14), are synthesized. Only the 24-mer is 5′ phosphorylated using standard methods. Equimolar amounts of the two oligonucleotides are mixed and allowed to anneal.

To produce cohesive ends for the adapter, DNA from plus and minus antibody immunoprecipitations is resuspended in 9 μl of *Nde*II[4] buffer (10 mM $MgCl_2$, 0.15 M NaCl, 10 mM Tris–HCl, pH 7.6) and digested for 1 h at 37°C with 4 U of *Nde*II (Boehringer). The reaction is stopped on ice and the DNA is precipitated with ethanol in the presence of 0.3 M sodium acetate, pH 5.2, and 20 μg glycogen as carrier. After precipitation, DNA (approximately 1 ng) is resuspended in 9 μl ligase buffer (10 mM $MgCl_2$, 5 mM β-mercaptoethanol, 2 mM ATP, 10 mM Tris–HCl, pH 7.5) containing linker at a final concentration of 0.5 μM. Ligation is carried out by the addition of 4 U of T4 DNA ligase (Boehringer) and incubation at 4°C for 24 h. We find that ligation can be a critical step: linker concentrations higher than 5 μM result in no PCR products. For 1 ng chromatin DNA 1 μmol linker is used. The ligated mixture is directly used as a template in a 100-μl PCR using *Taq* I polymerase, 250 μM each deoxynucleotide, and the corresponding reaction buffer. Under these conditions, the final $MgCl_2$ concentration is 2.5 mM. The primer used is the 20-mer oligonucleotide described above, added to a final concentration of 1 μM. Amplification is performed using one cycle of 94°C for 2 min; 35 cycles of 94°C for 1 min, 55°C for 1 min, and 72°C for 1 min; and 1 cycle of 94°C for 1 min, 55°C for 1 min, and 72°C for 10 min.

9. Southern Analysis and Mapping of Binding Sites in DNA

The immunoprecipitation enriches for DNA targets of a particular protein. This enriched chromatin fraction DNA can be used as a probe to scan genomic regions and determine the *in vivo* distribution of regulatory proteins. Typically, genomic clones covering the region of interest are digested with restriction enzymes, the DNA fragments are separated by agarose gel electrophoresis and immobilized onto a hybridization membrane, and the filter is hybridized with the immunopurified chromatin DNA fraction.

When the immunopurified DNA is used as a probe, the amplified DNA is freed from linkers prior to radiolabeling. After PCR (Section 8), samples are extracted once with phenol–chloroform, extracted once with chloroform–isoamylalcohol, and ethanol precipitated. Linkers are removed by digesting the DNA with 4 U of *Nde*II as described above and separating by gel filtration (Chroma-spin TE-100, Clontech). DNA probes (50–100 ng DNA) are routinely labeled by oligonucleotide random-primed DNA synthesis with [α-^{32}P]dCTP (sp act 3000 Ci/mmol, Amersham).

We apply a very stringent hybridization procedure (15), using Genescreen Plus nylon membranes (Du Pont). Briefly, membranes are prehybridized for 3 h at 65°C in 7% SDS, 1 mM Na-EDTA, 1% BSA, 0.5 M $NaHPO_4$, pH 7.2 (1 M $NaHPO_4$, pH 7.2 stock is 0.5 M Na_2HPO_4 with the pH adjusted with ca. 4 ml/liter *ortho*-phosphoric acid; at pH 7.2, $[Na^+] = 1$ M). Heat-denatured labeled DNA is added directly to the prehybridization solution, and hybridization is allowed to continue for 5–16 h at 65°C. Filters are washed at 65°C once for 10 min in 5% SDS, 1 mM Na-EDTA, 0.5% BSA, 40 mM $NaHPO_4$, pH 7.2, and at least four times for 5 min each in the same buffer but containing 1% SDS.

Figure 6 shows the results obtained using DNA from chromatin immunoprecipitated with anti-Polycomb antibodies as a probe against the BX-C genomic walk of *Drosophila.* Binding sites are scored all along the repressed region of the complex (*Ultrabithorax* and *abdominal-A* genes) whereas the active region (*Abdominal-B*) is devoid of Pc protein, consistent with the role of Pc as a chromatin domain-organizer (8). This result testifies to the substantial efficiency of the chromatin immunoprecipitation. In a conventional Southern hybridization analysis, the minimum amount of probe required for rapid (few hours) binding and detection (high specific activity) is at least 1 ng DNA.

10. Quantification of Hybridization Signals

In many cases the immunoprecipitated DNA will only hybridize to limited, discrete elements in a given genomic walk (e.g., if the immunoprecipitation utilized antibodies against a transcription factor). However, when analyzing chromatin components such as Polycomb protein, extended genomic regions may hybridize (see Fig. 6). In this case it could be useful to quantify the hybridization signal of each genomic fragment to determine the relative distribution of the immunoprecipitated protein on the DNA. Hybridization signals

[4] *Nde*II restricts DNA in a methylation-sensitive manner. Therefore this enzyme may not be appropriate for use in cell types where the genome is methylated.

are quantified using a PhosphorImager apparatus (Molecular Dynamics), and the integrated value of each band is determined using PC software (ImageQuant), connected to the PhosphorImager. The resulting values are normalized to take account of the molecular weight difference between bands, and these values can be plotted along the genomic walk, to give a profile of protein binding (see Fig. 7).

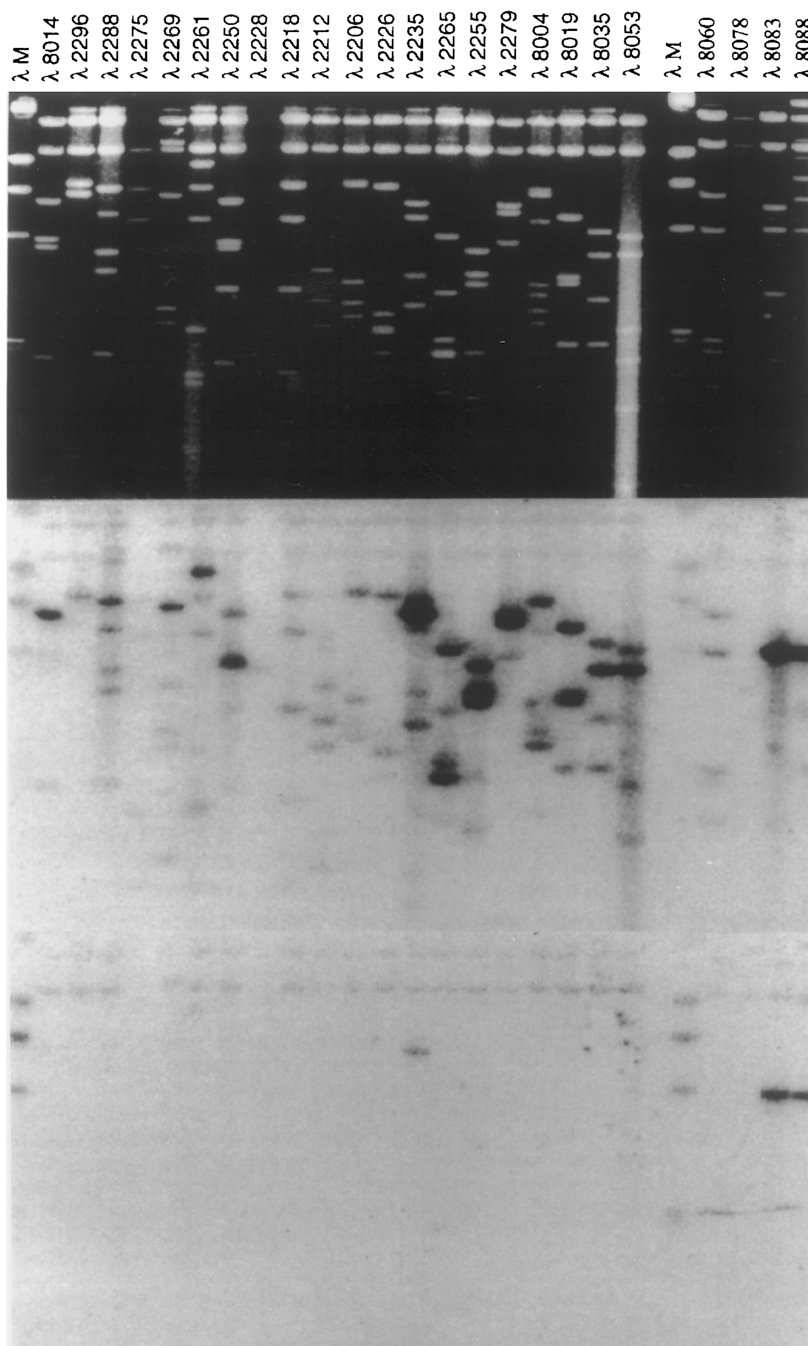

FIG. 6. Southern hybridization of immunoprecipitated and PCR-amplified fragments to the bithorax complex walk. (Top) Equal amounts (~0.5 μg) of each of 27 overlapping lambda Charon 4 clones covering the 300 kb of the BX-C walk were digested with *Eco*RI and resolved on a 0.5% agarose gel. λ2261 and λ2250 were digested with *Eco*RI–*Bam*HI and λ2235 was further digested with *Sal*I. The order on the gel from left to right reflects the proximal–distal order of the clones on the map. λ2255 and λ2265 were inverted by mistake. Visible molecular weight marker fragments (lambda *Hin*dIII) sizes are 23.1, 9.4, 6.6, 4.4, 2.3, 2.0, and 0.56 kb. (Middle) The gel shown above was transferred onto Gene Screen Plus membrane (Du Pont) by capillary blot and hybridized with ^{32}P-labeled total DNA from chromatin immunoprecipitation with anti-Pc antibodies. (Bottom) The same filter rehybridized with total DNA from a chromatin immunoprecipitation without antibodies. (From V. Orlando and R. Paro, *Cell* **75,** 1187–1198, 1993, reprinted with permission of Cell Press.)

One problem in such a quantification is to determine the level at which hybridization rises above background and is specific to the immunoprecipitation. For regions where the enrichment is large (for example in Polycomb immunoprecipitations), this background becomes negligible; however, the signal-to-background ratio may become significant in cases where a particular sequence is only weakly enriched during the immunoprecipitation. For example, if genomic DNA is hybridized for long periods to the genomic walk, a uniform hybridization of all bands may be seen. A number of solutions to this problem can be suggested. First, if sequences are known that do not interact with the protein of interest, then the amount of hybridization to these sequences can set the background level. Alternatively the signals generated from hybridizing DNA from control immunoprecipitations (labeled to the same specific activity) can be quantified and subtracted from the values of the actual immunoprecipitation.

It should also be noted that repetitive elements will always hybridize strongly (see bottom panel of Fig. 6 and white bars in Fig. 7). In these cases the signal resulting from specific immunoprecipitation cannot be accurately determined.

CONCLUSIONS

We have adapted this method to the analysis of low-abundance DNA-binding transcription factors. We have studied the distribution of the *Drosophila* transcriptional activators of the trithorax group, brahma and trithorax and of the homeodomain protein Abdominal-B (V. Orlando and R. Paro, unpublished results). We were able to map their relative distribution over the genomic region of the *empty spiracles* gene and found that a major binding site for all three proteins

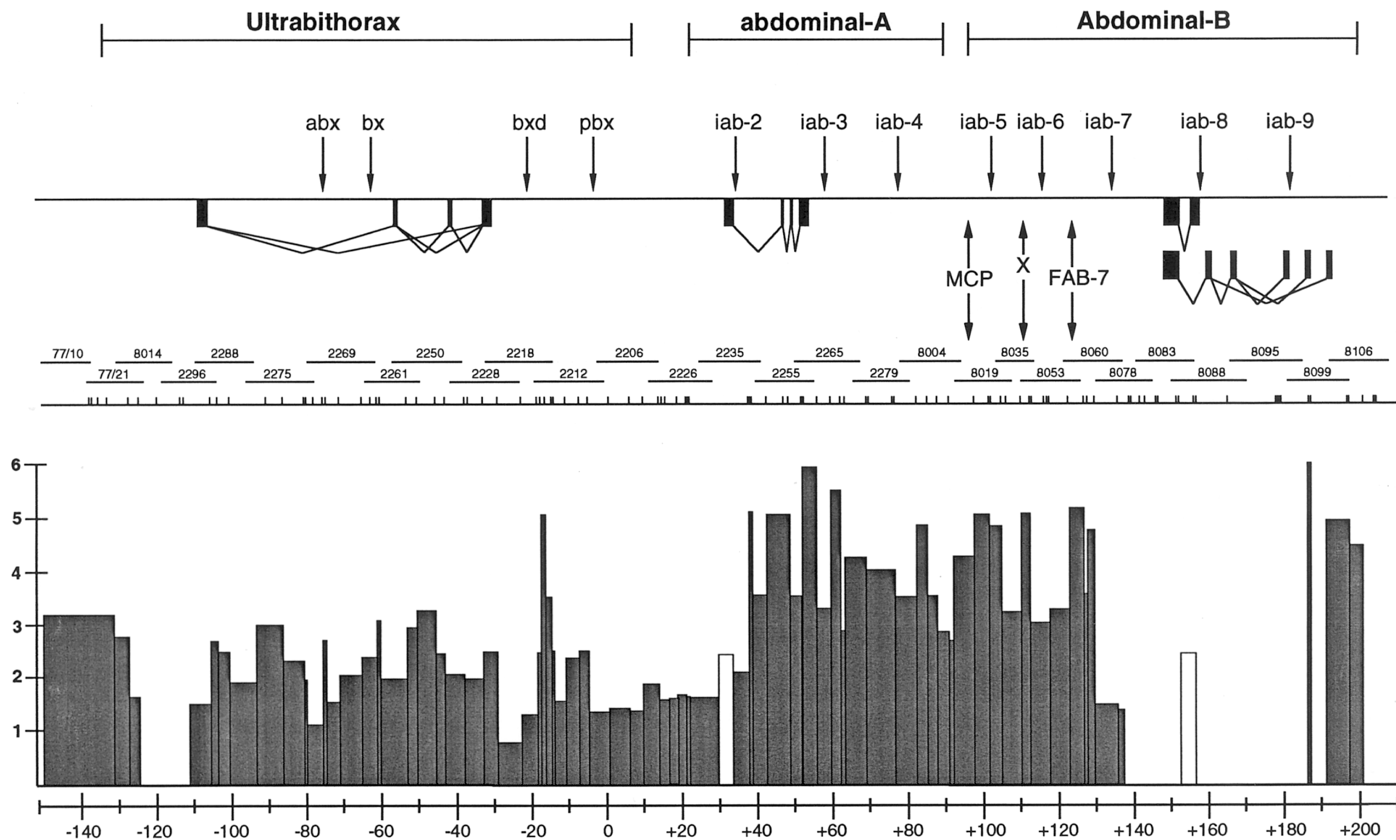

FIG. 7. Profile of Pc binding sites in the bithorax complex. The organization of the complex is depicted in the upper part of the figure. The regions of the *Ultrabithorax* (*Ubx*), *abdominal-A* (*abd-A*), and *Abdominal-B* (*Abd-B*) genes are drawn with their respective regulatory regions identified from genetic rearrangements. Transcription units are shown below with the corresponding exon structures as black boxes. The location of the presumptive boundary elements Mcp between iab-4 and iab-5, and Fab7 between iab-6 and iab-7 and of a postulated element X between iab-5 and iab-6 are shown. Below, the line depicts the distribution of *Eco*RI sites in the BX-C of Canton-S wild-type flies, and the λ clones covering the region. The bottom part of the figure shows the distribution of Pc in the complex as deduced from the quantitation of the hybridization intensities of the immunoprecipitated chromatin fragments. Bars represent in heights the absolute values of each individual fragment (given in arbitrary values) and in width the covered region on the map. The white bars indicate the fragments containing repetitive elements (M-repeats) that hybridized also with the control fraction without anti-Pc antibodies, thus making a precise quantitation impossible. (From V. Orlando and R. Paro, *Cell* **75,** 1187–1198, 1993, reprinted with permission of Cell Press.)

coincide with an enhancer element previously identified by reporter gene analysis (16). Interestingly, we find that the Abdominal-B protein *in vivo* has many binding sites scattered over an extended region of the *empty spiracle* gene. This distribution profile of a homeodomain protein obtained with the formaldehyde cross-linking method correlates with a similar analysis utilizing UV cross-linking (17; see also Biggin and Walter, 21). The sensitivity of the method relies heavily on the availability of high-quality antibodies that can immunoprecipitate the antigen of interest very efficiently. If this situation is given we find that the method also permits the detection of *in vivo* binding sites of low-abundance proteins. Indeed, this methodology has been utilized to isolate target genes of a mouse transcription factor, thus demonstrating that such an approach can also be applied to more complex genomes than *Drosophila* (18; H. Schöler, personal communication).

A major difficulty is presented in the discrimination of background versus real signal in the Southern analysis. In this respect, it is very helpful if known or suspected target sequences can be utilized in conjunction with nonspecific sequences as internal controls to establish the baseline level above which a hybridization signal can be taken as showing a real target element. In addition one should keep in mind that the hybridization profiles that are generated over extended genomic regions (see Fig. 7) reflect the fact that a three-dimensional, higher order structure was cross-linked. Thus, an immunoprecipitated fragment might not necessarily represent a directly interacting DNA partner of the protein but could also uncover a distant site which was in contact through the multimeric protein interactions of the chromatin structures. Therefore, the identification of interacting DNA elements by the *in vivo* cross-linking method will always require subsequent alternative assay systems to identify the physiologial function of the immunoprecipitated fragment in the context of the mechanisms of gene regulation, chromosome-mechanics, etc. Formaldehyde is one of the most frequently used fixatives in biological research as it penetrates the cell very efficiently and maintains the complex cellular structures in a remarkably intact form. The approach we present should provide a powerful tool to decipher the elusive higher order structures of chromatin.

REFERENCES

1. Varshavsky, A., and Ilyin, Y. V. (1974) *Biochem. Biophys. Acta* **340,** 207–217.
2. Moller, K., Rinke, J., Ross, A., Buddle, G., and Brimacombe, R. (1977) *Eur. J. Biochem.* **76,** 175–187.
3. Jackson, V. (1978) *Cell* **15,** 945–954.
4. Schouten, J. P. (1985) *J. Biol. Chem.* **260,** 9929–9935.
5. Jackson, V., and Chalkey, R. (1985) *Biochemistry* **24,** 6921–6930.
6. Solomon, M. J., and Varshavsky, A. (1985) *Proc. Natl. Acad. Sci. USA* **82,** 6470–6474.
7. Solomon, M. J., Larsen, P. L., and Varshavsky, A. (1988) *Cell* **53,** 937–947.
8. Orlando, V., and Paro, R. (1993) *Cell* **75,** 1187–1198.
9. Tomotsune, D., Shoji, H., Wakamatsu, Y., Kondoh, H., and Takhashi, N. (1993) *Nature* **365,** 69–72.
10. Varshavsky, A., Bakayev, V. V., Ilyin, Y. V., Bayev, A. A., Jr., and Georgiev, G. (1976) *Eur. J. Biochem.* **66,** 211–223.
11. Varshavsky, A., Sundin, O., and Bohn, M. (1979) *Cell* **16,** 453–466.
12. Harlow, E., and Lane, D. (1988) "Antibodies. A Laboratory Manual," Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY.
13. Zhao, K., Hart, C. M., and Laemmli, U. K. (1995) *Cell* **81,** 879–889.
14. Saunders, R. D. C., Glover, D. M., Ashburner, M., Siden-Kiamos, I., Louis, C., Monastirioti, M., Savakis, C., and Kafatos, F. (1989) *Nucleic Acids Res.* **17,** 9027–9037.
15. Church, M. G., and Gilbert, W. (1984) *Proc. Natl. Acad. Sci. USA* **81,** 1991–1995.
16. Jones, B., and McGinnis, W. (1993) *Genes Dev.* **7,** 229–240.
17. Walter, J., Dever, C. D., and Biggin, M. D. (1994) *Genes Dev.* **8,** 1678–1692.
18. Hess, H. (1996) Ph.D. thesis, University of Heidelberg.
19. McGhee, J. D., and von Hippel, P. H. (1975) *Biochemistry* **14,** 1281–1296.
20. McGhee, J. D., and von Hippel, P. H. (1975) *Biochemistry* **14,** 1297–1303.
21. Biggin, J., and Walter, M. D. (1997) *Methods Enzymol.* **11,** 215–224.