Alternative DNA-protein interactions in variable-length internucleosomal regions associated with *Drosophila Adh* distal promoter expression

Amy Ewel, Jeffrey R. Jackson and Cheeptip Benyajati*

Department of Biology, University of Rochester, Rochester, NY 14627, USA

Received December 12, 1989; Revised and Accepted March 1, 1990

ABSTRACT

Chromatin at the *Drosophila Adh* distal promoter displays an ordered but different conformation in different cell types as detected by a modified exonuclease protection assay and accessibility to endonucleases. In cells not transcribing *Adh* (ADH^-) sequences between -40 to $+30$ of the distal RNA initiation site exist as a DNA linker between positioned nucleosomes, and appear to interact with a specific DNA-binding protein. In contrast, a longer linker DNA, from -140 to $+30$, is bound in a multi-protein transcription initiation complex in cells that specifically transcribe the distal (adult) ADH RNA (ADH^+_A). These DNA-protein interactions can account for a localized open chromatin structure at the distal promoter in ADH^+_A cells. The observed mutually exclusive patterns of DNA-protein interactions in the linkers of different ADH cell types between -40 to $+30$ suggest a model for organizing alternative chromatin structure associated with gene regulation. Two DNA binding proteins, one being a TATA box binding factor, compete for overlapping sites to allow either assembly of a transcription initiation complex and transcription, or positioning of nucleosomes for stable repression.

INTRODUCTION

A single alcohol dehydrogenase gene (*Adh*) in *Drosophila melanogaster* encodes two mature ADH messages that are transcribed from two different, tandemly arranged promoters in a developmental- and tissue-specific manner (1, 2). A complex pattern of alternative promoter usage occurs throughout development and even within a single tissue (3, 4). How such intricate patterns of *Adh* transcription and repression are regulated is not well-understood. Recent studies indicate at least two levels of transcriptional control at the *Adh* locus, one involving changes in the concentration or activity of a sequence-specific DNA binding factor affecting the transcription initiation process (5), the other involving specific changes in chromatin structure of the promoter regions that may control the accessibility of cis-regulatory sequences to trans-acting factors and the general transcription machinery (6, 7).

Adf-1 is a sequence-specific DNA binding factor that binds upstream of both the *Adh* distal (adult) and proximal (larval) promoters but *in vitro* stimulates transcription initiation only from the distal promoter (8). The activity in stage-specific embryonic extracts correlates well with distal promoter activity during embryogenesis (5). Deletions of the Adf-1 binding site affected the distal promoter transcription in adult flies transformed with P-elements containing *Adh* gene constructs (9) and in transient expression assays in tissue culture cells (10). However, several observations suggest that Adf-1 may not be the only factor involved in regulating *Adh* expression. Adf-1 also binds *in vitro* to other promoters with different temporal transcription patterns during development (5, 49). Adf-1 is present in tissue culture cell lines that do not express the endogenous *Adh* gene (ADH^-), (8), and is functional in these cells in transient expression assays of exogenous *Adh* templates (6, 10, 11). Moreover, for the *Adh* gene the distal promoter TATA box and RNA initiation site are also necessary for *in vivo* expression both in adult flies and in tissue culture cells (9, 10). Recently, a new DNA binding factor has been identified that binds specifically to two sites flanking the Adf-1 site, the TATA box and the distal RNA initiation site (12, see Figure 4, of this study).

In vivo the distal and proximal promoter activities are associated with accessible chromatin structure. A switch from proximal to distal promoter expression (1, 3, 4) in ADH -expressing fat body cells of early and late third instar larvae is correlated with changes from DNase I hypersensitivity in the region of the proximal promoter and RNA start site to the distal promoter region (7). The nucleoprotein organization at the *Adh* locus in *Drosophila* tissue culture cell lines that show differential *Adh* expression has also been analyzed (6). Transcriptional activity specific to the distal promoter in ADH^+_A cells is correlated with a localized DNase I hypersensitivity around ~ -200 to $+50$ of the distal RNA start site. These same sequences are protected from DNase I in chromatin of all ADH^- cell lines studied. Such regions of nuclease hypersensitivity are thought to be nucleosome-free, and are probably accessible to non-histone proteins and to the transcription machinery (reviewed in 13, 14). This accessibility appears to be playing an important role in the regulation of *Adh* since it is clear that the ADH^- cell lines have the factors necessary to transcribe an exogenous *Adh* gene even though they

* To whom correspondence should be addressed

do not transcribe their endogenous copy (6, 10, 11).

In order to integrate the two levels of transcription control at the *Adh* distal promoter, the different chromatin conformations of the distal promoter (6) were analyzed for specific DNA-protein interactions in tissue culture cells using a high-resolution indirect end-labeling technique (15). Intact nuclei were digested with restriction endonucleases followed by exonucleases that can enter a multiprotein complex without major structure disruptions, as well as with the endonuclease DNase I that detects specific patterns of protein binding, and cuts at accessible sites. By analyzing the resulting DNA fragments on sequencing gels we were able to establish contact sites and boundaries of various proteins at or near single-nucleotide resolution. The results presented here demonstrate *in vivo* that the *Adh* distal promoter of ADH⁻ cells is not occupied by Adf-1 or the TATA binding factor, and that the binding sites for these factors are occupied instead by what appears to be a specifically positioned nucleosome and a novel factor Adf-2.

MATERIALS AND METHODS

Cells

Drosophila tissue culture cells were grown in M3 medium (43, 44), Schneider's *Drosophila* medium (GIBCO, for Schneider line 2, SL2) containing 10% fetal calf serum, and D-22 medium without serum (for K_{co}, obtained from Arno Greenleaf), at 23°C. The characterization of these cell lines into ADH⁻ and ADH⁺_A was described in (6). ADH transcripts are not detected in ADH⁻ cells whereas cells of ADH⁺_A lines transcribe their endogenous genes specifically from the distal *Adh* promoter at varying levels.

Modified exonuclease protection assay of chromatin

Nuclei were prepared from each *Drosophila* continuous cell line according to (16) with minor modifications. Cell pellets (usually 4 days old culture, from 6–8 T-75 flasks, 2–5 × 10⁸ cells) were resuspended in ice cold nuclear buffer (60 mM KCl, 15 mM NaCl, 5 mM MgCl₂, 0.1 mM EGTA, 15 mM Tris-HCl pH 7.4, 0.5 mM dithiothreitol (DTT), 0.1 mM phenylmethylsulfonyl fluoride (PMSF), 0.3 M sucrose), with 0.5% NP40, vortexed vigorously, and centrifuged through 1 volume of nuclear buffer + 1.7 M sucrose (stirred interface) at 10,500 rpm in an HB4, swinging bucket rotor. Nuclei were washed once with digestion buffer (nuclear buffer + 5% glycerol), and resuspended in the digestion buffer at approx. 5 × 10⁸ nuclei/ml. Restriction enzyme (300 and 1000 units/ml for approximately 10 to 50 µg DNA), RNase A (100 µg/ml), and endogenous exonuclease digestions were performed simultaneously for 1 hr at 37°C. Supplementary *E. coli* exonuclease III was added to the reaction during the last 15 min of incubation. SDS was then added to 0.5%, proteinase K to 100 µg/ml and the mixture was incubated overnight at 50°C. The DNA was extracted twice with 1:1 phenol/chloroform, twice with chloroform and precipitated with ethanol. For S1 nuclease trimming of single-stranded DNA tails, the reaction contained 4000 units/ml of S1 in 40 mM NaOAc, pH 5.2, 1 mM ZnSO₄ and 0.25 M NaCl at 30°C for 15 min. The DNA was repurified by organic extraction and ethanol precipitation.

Digestion of nuclei with DNase I

Nuclei were isolated as described above with the following exceptions. Digestion buffer contained in addition 0.5 mM CaCl₂, and no glycerol or RNase A. Nuclei were suspended in

digestion buffer at an OD₅₅₀ of 10 (estimated as a 1:100 dilution). DNase I was added at concentrations ranging from 0 to 40 U/ml, digestion was carried out for 3 min at 25°C, and stopped by addition of EDTA to 12.5 mM, SDS to 0.5%, and proteinase K to 200 µg/ml, and incubated at 50°C overnight. The DNA was extracted twice with 1:1 phenol/chloroform, and once with ether. RNase A was then added to 50 µg/ml, and the mixture incubated at 37°C for 2 hr. SDS and proteinase K were added again to 0.5% and 200 µg/ml respectively, and incubated at 50°C for 4 hr, followed by 2 phenol/chloroform extractions, an ether extraction, and ethanol precipitation.

The indirect end-labeling technique: restriction digestion, gel electrophoresis and hybridization

Low resolution, double-stranded DNA analysis

DNA samples (10 µg) were digested with appropriate restriction enzymes and separated on 25-cm long agarose gels made in TBE. These secondary restriction enzyme digestions were a. PstI (–3014A and +5000A) b. EcoRI (–1320A and +3350A) and c. EcoRI (–1320A)/BamHI (+1257A) double digests. The internal size standards used on each gel were: ϕX174 DNA cleaved with HaeIII, λ DNA cleaved with HindIII, and purified *Drosophila* genomic DNA (from cell lines S2 and 1006-2) cleaved with appropriate restriction enzymes near the genomic locations of interest. After blotting onto nitrocellulose, the filters were probed with ³²P-labeled, nick-translated DNA (specific activity > 1 × 10⁹ cpm/µg DNA). The ³²P-labeled probes about the PstI (PstI, –3014A to HindIII, –1848A), the EcoRI (EcoRI, –1320A to XbaI, –663A) and the BamHI (BamHI, +1257A to BanII, +719A) sites respectively.

High resolution, single-stranded DNA analysis

Analysis of cleavage patterns at single-nucleotide resolution was carried out using the methods originally developed by (15), as described in (36). 12.5 µg samples of genomic DNA were digested to completion by XmnI, phenol/chloroform extracted once, ethanol precipitated, and resuspended in 3 µl of loading buffer (90% formamide, 1 × TBE, 0.02% bromophenol blue, 0.02% xylene cyanol). DNA fragments were denatured by heating in boiling water, and separated on standard 6% acrylamide 40 cm sequencing gels. The DNA was then electroblotted onto Nytran membrane (Schleicher and Schuell), and UV-crosslinked to the membrane (dry) with a dose of 160 J/m². The filters were then hybridized with a short riboprobe, which abuts the reference end, in 10 ml of hybridization solution as described by (15), and modified by (45) for 18 hr at 58°C. The riboprobes was transcribed from an XmnI (–248A) to FspI (–130A) fragment cloned into the 'Bluescript KS' plasmid (Stratagene), using T7 or T3 RNA polymerases. The fragments were cloned into the SmaI site and the plasmid was linearized with either PstI (T7), or BamHI (T3) prior to transcription. Reaction conditions were as described by (36). Following hybridization the filters were washed at 58°C as described by (15), and exposed with an intensifying screen for 4 to 10 days.

Note that the bands in the G+A sequencing ladder run 1.5 nucleotides ahead of the corresponding bands in the DNase I digested DNA. This is because chemical sequencing eliminates a nucleoside (46), causing a displacement of one nucleotide, and it leaves a 3' phosphate so that when mapping from the 5' end there is an additional displacement of a half nucleotide (47).

RESULTS

Probing chromatin structure with restriction endonucleases

Several restriction endonucleases were used to cut chromatin of the *Adh* locus to survey appropriate entry sites for exonucleases. Table 1 compares the accessibility measurements of selected cleavage sites in nuclei of several ADH⁻ and ADH⁺_A cell lines. To ensure that limit digestions were achieved under our experimental conditions, 2 nuclease concentrations were employed. All enzymes gave partial chromatin cleavages at all detectable sites at limit digestions, nevertheless the percentage of cleavage at each particular site, as measured by quantitative Southern blot hybridizations, is reproducible for each cell line. These results agreed in general with those obtained by DNase I probing (6), and they indicate that the distal promoter chromatin in the ADH⁺_A cell type is in an open conformation. The accessible region specific to the active distal promoter appears to be localized to between ~ -202/-168 and +51, coinciding with a defined DNase I hypersensitive site present only in ADH⁺_A cells. The ADH protein coding and the 3' untranslated sequences appear to be equally protected from several restriction enzymes and DNase I in chromatin of both cell types (Table 1, also data not shown, and 6), consistent with previous detections of an array of positioned nucleosomes over these regions in both cell types (6, 7).

Endogenous exonuclease activity similar to *E. coli* exonuclease III (exo III)

We analyzed DNA-protein interactions *in situ* in chromatin by an exonuclease protection assay and the indirect end-labeling detection technique (16). Typically, exonuclease III (exo III) is used to digest double stranded DNA processively in a 3' to 5' direction from the free 3'OH termini generated by restriction enzymes until the enzyme is blocked by specific protein binding or features of the DNA structure in chromatin that cause termination are encountered. However, as first discovered by Wu, (16), we have confirmed that all *Drosophila* cell lines studied possess an endogenous exonuclease activity. Under our experimental conditions (1 hr, 37° C incubation with a restriction endonuclease, endogenous exonuclease, and RNase A, see below) this activity does not affect the mobility of double-stranded DNA resulting from limit restriction digestions of chromatin unless the DNA is subsequently treated with S1 nuclease (see Figure 7A). Both endogenous exonuclease and exo III appear to recognize the same features of nucleoprotein structure to generate the same patterns of resistant sites in chromatin (see below).

We have also observed that the presence of RNase A during the exonuclease protection assays of chromatin, while not affecting the cleavages by restriction endonucleases and DNase I, appears to greatly enhance the activity of endogenous exonuclease as measured by the ability to penetrate chromatin from an entry site. The basis for these observations is not known. However, the specificity of the exonuclease for DNA-protein structure appears unchanged based on experiments with or without RNase A (data not shown).

Barriers against exonuclease digestion

Figure 1 shows low-resolution analyses of the differences in exonuclease digestion patterns of the distal promoter chromatin between ADH⁻ and ADH⁺_A cell types. SalI sites at -63 and -69 provided the entry sites for endogenous exonuclease and exo III added during the last 15 min of incubation. SalI cuts at

Figure 1. Chromatin exonuclease digestion patterns in the distal promoter region viewed 5' and 3' from the -63A, -69A SalI entry sites. Exonuclease barriers were mapped in nuclei prepared from an ADH⁻ (5697-70) and an ADH⁺_A (1006-2) cell line by digesting with 1000 units/ml of SalI (to generate entry sites for exonucleases), RNase A (100 µg/ml), and endogenous exonuclease for 1 hr at 37°C. During the last 15 min of incubation, exo III was supplemented as indicated. DNA was isolated, digested with S1 nuclease to trim the single-stranded tails. S1 nuclease-treated DNA was then double-digested with EcoRI (-1320A) and BamHI (+1257A), fractionated on a 1% agarose gel (10 µg DNA/lane), blotted and hybridized with the 657 bp EcoRI-XbaI (-1320A to -663A), ³²P-labeled probe (Top panel). The filter was then stripped and rehybridized (Bottom panel) with the 538 bp BamHI-BanII (+1257A to +719A) probe. The positions of the genomic size markers are shown. The positions of ϕ X174 DNA markers cleaved with HaeIII are indicated as dots. α and T, V, W mark the chromatin exonuclease resistant sites upstream, and I and C, D, E, F mark the sites downstream of the -63A, -69A SalI cleavages in ADH⁻ and ADH⁺_A cells respectively.

the -63 and -69 sites at a very low level in ADH⁻ cells, and at a very high level in ADH⁺_A cells (see Table 1). Consistent with the SalI sites in ADH⁻ chromatin being within the endonuclease-protected (see Table 1, refs 6, 7) and exonuclease impenetrable region (Figure 1, top and bottom), a large portion of the intact EcoRI-SalI and BamHI-SalI restricted DNA bands remained at limit exonuclease digestion (Figure 1, top and bottom panels, lanes 2-4). In contrast, most of the EcoRI-SalI band could be progressively digested to generate multiple exonuclease

Table 1. Comparison of chromatin accessibility to restriction endonucleases limit digestions between ADH⁻ and ADH⁺_A cells. Nuclei from several ADH⁻ and ADH⁺_A cell lines (approximately 50 µg of DNA), about 5-day-old in culture, were digested with various restriction enzymes at 300 and 1000 units/ml for 1 hr at 37°C. The positions of restriction endonuclease cleavage in chromatin were determined by a secondary restriction digestion of 10 µg of purified genomic DNA, gel electrophoresis, Southern blot hybridization and the indirect end-labeling technique as described in Materials and Methods (see Figure 7A for an example autoradiogram). Quantitation of the restriction enzymes-limit digests in each gel lane was performed by densitometric tracings of the autoradiograms. Note that all the restriction sites reported here appeared as visible bands on the autoradiograms at limit digestions, which were usually achieved at 300 units/ml of restriction nucleases. Averaged limit-digest values for the 2 enzyme concentrations are shown for selected cleavage sites around the *Adh* distal promoter and protein coding sequence. The lowest limit of detection was about 0.1% cleavage. The measurements may be somewhat underestimated if the cleavage sites are multiple, but they did not affect the outcome of chromatin comparison between ADH⁻ and ADH⁺_A cells. The distances in bp from the distal (adult, A) RNA start site are shown. The distal TATA box is located at -26A to -32A. ATG, initiation and TAA, termination codons for the ADH polypeptide are at positions +778A and +1681A respectively. The BanII site at -168A is polymorphic and is present in only two of the cell lines studied.

Cells	% Cleavage in Chromatin										
	XmnI -248A	DdeI -202A	BanII -168A	HhaI -129A	SalI -69A/-63A	AvaII -14A	DdeI +9A	ScaI +51A	XmnI +222A	BamHI +1257A	ScaI +2628A
ADH⁻											
2588-7	0.1	0.5	—	—	1	—	0.5	26	1	22	1
PC1-8	0.1	—	—	—	1	—	—	20	1	—	1
2588-6	0.1	—	—	—	1	—	—	23	1	—	1
SL2	0.1	1	8	1	1	1	1	26	1	—	1
ADH⁺_A											
OX736-3	0.5	2	23	15	46	10	14	47	1	—	1
UC13-1	0.5	2	—	23	35	—	28	57	1	20	1
OX736-4	0.5	—	—	—	40	—	—	56	1	—	1
1006-2	0.5	—	—	10	18	14	—	40	1	—	1

barriers in ADH⁺_A chromatin (Figure 1, W, V, T, top panel probing from the 5' direction), also when probed from the 3' direction the BamHI-SalI band behaves similarly (Figure 1, C, D, E, F, bottom panel). Regions T and F thus mark the 5' and 3' borders in the ADH⁺_A distal promoter chromatin where exonucleases cannot penetrate further (impenetrable barriers).

Two further points can be made about the results shown in Figure 1. First, the endogenous exonuclease appears to be quite active under these experimental conditions; in some preparations limit digestions were attained without supplementary exo III. These observations are not ADH-cell-type or cell-line-specific, and were possibly due to minor variations in nuclei preparations. Second, although our results suggest the precursor-product relationship between the barriers closer to the entry sites and the ones further away (Figures 1, and 7B with the ScaI site at +51 as the entry site, and unpublished results for other ADH⁻ and ADH⁺_A cell lines), cell-type-specific multiple exonuclease barriers were reproducibly observed at limit digestions (see below). These observations would suggest a possible heterogeneity in the cell population, or an unusual DNA structure in chromatin that causes exonucleases to terminate in the nuclei population (see below and Discussion). Nevertheless, the patterns of these barriers are not random; they are highly specific to either the ADH⁻ or ADH⁺_A cell type and correlate with the distal promoter transcriptional activity.

High-resolution analysis of the exonuclease barriers

Figures 2 and 3 display the single-stranded exonuclease resistant DNA fragments, generated by using the +51 ScaI entry site and digesting with exonucleases as described in Figure 1, separated on sequencing gels, and analyzed by the high-resolution indirect end-labeling technique (15). Single-base resolution was obtained from -170 to -50 in Figure 2 (sense strand), and from -95 to +18 in Figure 3 (antisense strand). It is evident that the resistant sites are ADH cell-type-specific and very reproducible; these are summarized aligned with the DNA sequences in Figure 4. The 5' impenetrable site in ADH⁻ chromatin is located at

-66 (the last nucleotide detected was -65), whereas exonuclease resistant barriers extend to -165 in ADH⁺_A chromatin.

Several observations are of note. First, endogenous exonuclease appears to digest double-stranded DNA in chromatin processively in a 3' to 5' direction as indicated by analysis of exonuclease-digested samples without S1 nuclease treatment and strand-specific probing (data not shown). The enzyme appears to show no sequence specificity at the resistant sites (see Figure 4) indicating that the resistant sites observed in chromatin of both cell types are due to DNA-protein structure and their interactions.

Second, there are both similarities and differences in the patterns of resistant sites in the region -36 to +32 between the ADH⁻ and ADH⁺_A cell types. Strong and highly defined (1, 2 or 3 base wide) resistant sites occur periodically, several at about 10-base spacings, which suggest that these DNA regions in both cell types are structured in a non-random fashion to cause exonuclease pausing at intervals corresponding to the periodicity in the DNA double-helix. Superimposed on this periodicity are two mutually exclusive resistant sites: one is present only in ADH⁻ cells at +9 and +11, within an *in vitro* binding site of a sequence-specific DNA binding factor, Adf-2 (12), the other site is present only in ADH⁺_A cells at -25 and -26, within the TATA box, the likely binding site of the TATA binding factor (20, 21). Other differences in resistant sites between cell types exist throughout this -36 to +32 structured region as indicated by the double-headed arrows in Figure 3, and summarized in Figure 4. Except for a possible sequence polymorphism near the distal RNA start site (48), it is likely that these differences in resistance patterns between the cell types are due to alternative DNA-protein interactions in this region.

Third, the exonuclease resistant sites specific to ADH⁺_A cells, from -48 to -89 (region X, Figures 2 and 7B), coincide with the *in vitro* binding site of a sequence-specific positive transcription factor, Adf-1 (8). No factor has yet been identified *in vitro* for the DNA-protein interactions in region W (Figures 1, 2, and 7B), from -99 to about -140.

Fourth, although differences in chromatin between ADH⁻

Figure 2. High resolution analysis of exonuclease resistant sites from the +51A ScaI cut in the *Adh* distal promoter chromatin. S1 nuclease-treated genomic DNA samples, generated in the modified exonuclease protection assay in chromatin (nuclei digested with ScaI, RNase A, endogenous exonuclease, supplemented with 5000 units/ml exo III) were digested with XmnI, and analyzed by the high-resolution indirect end-labeling technique on a 6% sequencing gel using a 32 P-labeled riboprobe complementary to the sense strand, spanning the -248A XmnI (-243A is the actual nucleotide reference end) and the -130 FspI sites. Lane 1: single-stranded genomic size markers generated by restriction cleavages run on the same gel as the rest of the samples. Lane 2: G+A genomic sequence markers. Lanes 3-4, and 7-8 (a longer exposure of 3-4): chromatin prepared from 2 different ADH^{-} cell lines (SL2, K_{co}). Lanes 5-6, and 9-10 (a longer exposure of 5-6): chromatin prepared from 2 different ADH^{+} cell lines (OX736-3, 1006-2). Exonuclease resistant regions and their boundaries described in Figures 1 and 7B are indicated. The extended open boxes indicate the regions not clearly observed at lower resolution.

Figure 3. Comparison of exonuclease resistant sites between -80 and +50. S1 nuclease-treated genomic DNA samples, generated in the modified exonuclease protection assay in chromatin as described in the legend to Figure 2, were digested with XmnI and analyzed by the high-resolution indirect end-labeling technique using a 32 P-labeled riboprobe complementary to the antisense strand spanning the -248A XmnI (-243A is the actual nucleotide reference end) and the -130 FspI sites. Lane 1: single-stranded genomic size markers generated by restriction cleavages. Lanes 6-9 are a longer exposure of lanes 2-5. Lanes 2 and 3: chromatin prepared from an ADH^{-} cell line, 1704-2, without (2) and with (3) supplementary 5000 units/ml exo III. Lanes 4 and 5: chromatin prepared from an ADH^{+} cell line, 1006-2, without (4) and with (5) supplementary 5000 units/ml exo III. Double-headed arrows indicate the sites where differences between exonuclease resistance in chromatin of the ADH^{-} and ADH^{+} cell types are reproducibly observed. The double-headed dashed arrow at -1 points to a possible sequence polymorphism between the different alleles of *Adh* gene (48).

Figure 4. Summary of specific barriers to exonucleases aligned with nucleotide sequences of the *Adh* distal promoter. Data for the ADH^+ cell type are summarized above the top (sense) strand, while those for the ADH^- cell type are below the bottom (antisense) strand. The endogenous exonuclease appears to digest double-stranded DNA in chromatin processively in a 3' to 5' direction from the sites generated by restriction endonucleases. The DNA samples were then digested with S1 nuclease to trim the single stranded tails before being analyzed on the sequencing gels. The RNA start site is at +1. The TATA box (-32 to -26), the *in vitro* binding site for a transcription factor Adf-1 (-86 to -46), and the 2 ten-bp direct repeats that bind Adf-2, a sequence-specific DNA-binding factor, are marked between DNA strands. Striped lines mark several restriction endonuclease cut sites; the numbers indicate the extent of chromatin cleavage in ADH^- and ADH^+ cells (from Table 1). Vertical bars locate exonuclease resistant sites; their heights indicate relative resistance at that barrier. The stippled boxes indicate DNA sequences where exonucleases cannot penetrate in chromatin at limit-digestion. The open box indicates the resistant regions that supplementary exo III can enter at high concentrations.

and ADH^+ cells can be clearly seen 5' of $\sim +53$ (as described above), the region 3' of $\sim +53$ in both cell types when probed by exonucleases are remarkably similar; only impenetrable barriers are observed (data not shown, see summary in figure 4).

Impenetrable barriers are probably nucleosome cores

Several studies using exo III to digest isolated bulk mono- and di-nucleosomes and reconstituted mononucleosomes showed that the exonuclease is strongly impeded by the border of the histone octamer in the nucleosome core (for examples, see 17–19). Our analysis of the *ADH* protein coding region by the modified exonuclease protection assay also indicated that a previously identified positioned nucleosome at the +1257 BamHI site in both *ADH* cell types (6, 7) impeded the digestion by endogenous exonuclease (data not shown). Thus it appears likely that the impenetrable barriers identified above are nucleosome cores, which seem to be specifically positioned with respect to DNA sequences in both *ADH* cell types, in a cell-type-specific manner in the majority of cells in the populations. Interestingly the 3' barrier is the same in both types of cells, but the 5' barriers are

at different locations indicating that there are different length linkers between positioned nucleosomes in ADH^- and ADH^+ cells (Figure 4).

Genomic footprinting by DNase I at the *Adh* distal promoter

A corroboration for the ordered chromatin structure around the *Adh* distal promoter and RNA initiation site, with putative positioned nucleosomes and specific non-histone protein binding regions, was obtained by DNase I digestion of nuclei. Endonucleases like DNase I have limited access to DNA bound to a surface such that only one side of the DNA helix is accessible for cleavage (22). The only sites exposed to the enzyme are on the far side from the surface, and recur with the periodicity of the DNA double helix. Known protein-DNA interactions that give rise to these 10-base periodic patterns of DNase I cleavage are a 146 bp-DNA wrapped around a histone octamer (23–26), DNA bound to bacterial DNA gyrase (27), DNA bound to the mammalian and *Drosophila* TATA binding factors (20, 28), and DNA constrained into a small loop structure by 2 interacting λ repressor molecules bound at 2 separate sites (29).

Figure 5 shows the genomic footprinting experiments to

Figure 5. Genomic footprinting by DNase I in the region surrounding, and downstream from the distal RNA start site. The reference end of these single-stranded DNA fragments was created with XmnI, which has a site at -243 . The blots were hybridized with riboprobes transcribed from the plasmid containing the -243 to -130 region. *Panel A* shows cleavage in chromatin on the sense strand. Lane 1: purified genomic DNA digested with DNase I to an average double-stranded fragment size of 1.5 kb. Lanes 2–5: chromatin isolated from K₅₆₂ cells (ADH⁻) without DNase I digestion or digested to an average size of >23 , ~ 23 , and ~ 20 kb respectively. Lanes 6–8: chromatin isolated from 1006-2 cells (ADH⁺) without DNase I digestion or digested to an average size of ~ 23 , and ~ 10 kb. *Panel B* shows cleavage in chromatin on the antisense strand. Lanes 3–7: chromatin isolated from S2 cells (ADH⁻) without DNase I digestion or digested to an average size of >23 , ~ 23 , ~ 20 , and ~ 15 kb. Lanes 8–12: chromatin isolated from 1006-2 cells (ADH⁺) without DNase I digestion or digested to an average size of $>>23$, >23 , ~ 23 , ~ 20 kb. Lanes 13–14: purified genomic DNA digested with DNase I to an average size of ~ 10 , and ~ 2 kb. M: genomic size markers generated by restriction cleavages. G+A: genomic sequence markers. The cross-hatched boxes indicate the positions of the TATA box, and previously defined *in vitro* factor binding sites, the arrows point out the pattern of DNase I cleavage, and the brackets mark the regions in ADH⁺ cells that lack any pattern, and appear specifically nuclease-protected. Open circles indicate positions of specific protection from cleavage, and solid circles indicate specific enhanced cleavage positions.

Figure 6. Schematic representation of the comparative DNA protein-interactions at the *Adh* distal promoter in chromatin of ADH^- and ADH^+_A cell types. An interpretation of the combined data from restriction accessibility, exonuclease protection, and specific DNase I cleavage (arrows) and DNase I protection (open bars) experiments is shown for the distal promoter complex in ADH^+_A (Top), and ADH^- cells (Bottom). Locations and boundaries of specific DNA-protein interactions are indicated aligned to the distance from the distal RNA initiation site, assigned as in Discussion. Putative nucleosome cores are shaded for their apparently different DNA-octamer interactions. 1 represents a specific exonuclease resistant region seen only in the ADH^+_A cell type.

investigate chromatin in the vicinity of the 3' putative positioned nucleosome. The data for both the sense (A) and antisense (B) strands are presented. These data are also summarized and interpreted together with the restriction enzyme accessibility and the exonuclease protection data in Figure 6. In ADH^- cells there is a strong, well-defined 10-base periodicity of DNase I cleavage on both DNA strands, staggered by a few bases from each other, downstream of the RNA start site (Figure 5). The pattern of periodic cleavage is different from the naked DNA pattern downstream of +65. The pattern extends for at least 14 cleavage sites, is generated early during DNase I digestion, and appears to be stable over a range of enzyme concentrations. Taken together our results strongly support the notion that a precisely positioned nucleosome core is present at the 3' exonuclease impenetrable barrier in ADH^- chromatin with a border at or near +30, where the changes in digestion patterns of nucleases occur (see Figure 6 and Discussion). Immediately 5' of this boundary protected and enhanced ADH^- chromatin-specific cleavage sites are detected on the antisense strand in the *in vitro* binding site of Adf-2.

In ADH^+_A nuclei the 3' exonuclease impenetrable barrier is

very similar to that in ADH^- nuclei (described above, see Figure 4) and a similar DNase I cleavage pattern to that in ADH^- cells is also observed near the RNA start site on both DNA strands to around +37 on the sense, and around +65 on the antisense strand, then it is disrupted on both strands for at least 100 bp (Figure 5). It appears that ADH^+_A cells have a nucleosome positioned at this region with a 5' border at or near +30 (Figure 6), but it is structurally altered by the transcription process. Other investigators have similarly observed that arrays of nucleosomes are disrupted in active transcription units (7, 13, 14).

Although the restriction enzyme accessibility (Table 1), and exonuclease protection data (Figures 1, 2, and 3) suggest that a positioned nucleosome core may be located upstream of -140 in ADH^+_A , and -60 in ADH^- chromatin, attempts to corroborate this with DNase I digestions have been inconclusive. While different patterns can be seen there is not a clear cut 10-base ladder as seen at the 3' putative positioned nucleosome. The digestion patterns do however suggest that there may be different DNA-protein interactions at this region in the two cell types.

Dynamic interactions responsible for the open chromatin conformation at the active distal promoter

Apart from the differences in the distal promoter chromatin between cell types, variations within a cell type were also seen. Analysis of the 1006-2, ADH⁺_A cell line during growth in culture indicated that DNA-protein interactions within the active distal promoter are dynamic. Figure 7A shows that the accessible chromatin domain specific to the active distal promoter (~ -202 to +51, see Table 1) can undergo a localized structural change detected as a gradual two-fold increase in accessibility, as cultures age, at restriction sites within the domain while other cleavage sites outside the chromatin domain are not affected (e.g. the nearby XmnI sites). Parallel analysis of DNA-protein interactions by the modified exonuclease assay indicates quantitative changes in specific exonuclease barriers in older cultures (Figure 7B). Region X (the Adf-1 binding site), and region W are clearly under-represented at all stages of exonucleases digestions, while barriers Y and Z (the TATA factor binding site), and the 5' putative positioned nucleosome barriers V and T appear unchanged. Thus changes in specific DNA-non histone protein interactions at the distal promoter appear to influence the local chromatin structure. Of the 3 identified DNA binding sites in ADH⁺_A chromatin, the TATA binding interactions appear to remain stable and strong during the cell growth cycle in culture.

DISCUSSION

We reported a detailed chromatin organization and associated DNA-protein interactions in the distal promoter of the *Drosophila Adh* gene. We have provided evidence for the *in vivo* cell-type-specific binding of protein factors to the DNA motifs previously defined *in vitro*, and for the involvement of chromatin structure in regulating gene expression at the distal promoter. In ADH⁻ cells it appears that while they are functionally present, positive factors Adf-1 and the TATA binding factor are not bound to the distal promoter. Instead the binding sites for these factors are occupied by a specifically positioned nucleosome and Adf-2.

Nature of DNA-protein interactions: the active distal promoter

Occupation of the TATA box, the Adf-1 binding site and region W is ADH⁺_A chromatin specific. The protein factors appear to interact with DNA as a complex on the same DNA molecule; we could detect the exonuclease barriers further away from the restriction entry sites only after the appearance and at the expense of the ones closer (Figures 1 and 7B). It appears that this multi-protein complex is involved in specific transcription initiation of the distal promoter *in vivo* as discussed below.

The interactions at the distal TATA box and surrounding nucleotides from -36 to -22, and the downstream sequences to +32 in ADH⁺_A cells are consistent with the binding of the TATA transcription factor previously characterized *in vivo* (16), and *in vitro* (20, 21). A partially purified TATA binding activity from *Drosophila* tissue culture cells and embryos stimulates transcription initiation *in vitro*, and interacts somewhat differently with the TATA box of different promoters. The protected nucleotides at the promoters of several genes assayed by DNase I footprinting usually extend from -40 to +25/+35, which include the TATA box, the RNA start site, and a portion of the leader sequence (20). Competition experiments indicated that common factors are involved in producing this large bound complex, but the binding requires only the TATA box and not any specific sequences downstream (21). Similar observations

have been noted for TFIID, the TATA binding factor partially purified from HeLa cells (28, 30). For both *Drosophila* and mammalian activities unusual DNase I footprints were observed on some promoters; the TATA box is uniformly protected on both strands while the downstream region displays a 10 base periodic cleavage pattern, staggered by a few bases on the 2 strands (20, 28). This pattern of TFIID contacts downstream of the RNA start site appears to be important for the *in vitro* transcriptional activities of several promoters (31). A similar pattern of periodic interaction is also apparent in the region downstream of the *Adh* distal TATA box detectable by both exonucleases and DNase I digestions of ADH⁺_A chromatin, but not in ADH⁻ chromatin even though TATA binding activity is present (Figures 2, 3, 5, ref. 20).

Protein-DNA contacts between -44 to -89 unique to ADH⁺_A chromatin show a distinct pattern from the regions immediately 5' and 3'. Highly accessible restriction sites (-63/-69 SalI), exonuclease resistant sites as well as DNase I hypersensitive and protected regions are clustered from -48 to -68 whereas immediately 5' the exonuclease sites are more dispersed. Together the 2 regions match the Adf-1 *in vitro* binding site (-46 to -86). First identified in an ADH⁻ cell line K_{co}, Adf-1 is a sequence specific DNA binding factor that binds to both the *Adh* distal and proximal promoters as well as other regulated promoters, e.g. antennapedia P1 and dopa decarboxylase (5, 8). The Adf-1 binding site is required for *Adh* distal promoter transcription initiation *in vitro* (8) and *in vivo* (9, 10). Our results here indicate that while present in both ADH⁻ and ADH⁺_A cells, Adf-1 appears to occupy the distal promoter only in ADH⁺_A cells (Figures 2, 3).

Region W interactions (-98 to -143) are detected in all ADH⁺_A nuclei and represent quite strong and stable exonuclease barriers (Figures 1, 2). No functional activity nor DNA binding factor has yet been assigned, although altered interactions in this region and the Adf-1 site appear to affect the structure of the open chromatin domain (see Figure 7). Due to its high accessibility to the restriction cleavage when compared to ADH⁻ chromatin, it is unlikely that the DNA is bound to a nucleosome core (see Table 1).

The inactive distal promoter

Although both Adf-1 and the TATA binding factor are present and with demonstrated functional activity in ADH⁻ cell lines (6, 8, 10, 20), they clearly do not bind at the *Adh* distal promoter. Instead, a positioned nucleosome is located at the Adf-1 site and a new protein-DNA contact is observed at the *in vitro* binding site of Adf-2 within the TATA binding domain. Adf-2 is a sequence-specific DNA binding factor present in nuclear extracts of all *Drosophila* cell lines studied (12). DNase I footprinting and methylation interference assays identified 2 specific binding sites within the *Adh* locus involving ten-bp direct repeats (Figure 4), between -189 to -220 on the sense, and -187 to -211 on the antisense strand (Adf-2a), and between -7 to +18 on the sense, and -7 to +21 on the antisense strand (Adf-2b). The observed ADH⁻ chromatin-specific DNase I hypersensitive and exonuclease resistant nucleotides at the Adf-2b site reported here are crucial for Adf-2 binding *in vitro* (12). Work in progress suggests that the Adf-2a site may also be occupied in ADH⁻ nuclei. Although nearly 200 bp apart, the 2 Adf-2 binding sites and bound proteins could be closer together in nuclei as the DNA in between is apparently bound to a positioned nucleosome core in ADH⁻ chromatin.

Our results indicate that *in vivo* the distal promoter is organized with precise positioning of non-histone proteins and nucleosome cores into either an active or inactive conformation for transcription initiation. The patterns of DNA-non-histone protein interactions are mutually exclusive, apparently involving different sequence-specific DNA binding factors with overlapping DNA contact sites. Based on these observations we suggest a working model for organizing alternative chromatin structure in which 2 sequence-specific DNA binding proteins, the TATA box binding factor and Adf-2 compete for their overlapping binding sites, perhaps during DNA replication and chromatin reassembly (34, 35). Stable binding of the TATA box binding factor and further contact with the region 3' of the RNA initiation site, binding of Adf-1 and interactions in region W could allow assembly of a stable transcription initiation complex that excludes Adf-2 from binding and a nucleosome core from assembling at the distal promoter. This chromatin structure is thus accessible to the general transcription machinery. In contrast if stable Adf-2

interactions were established at the Adf-2 sites, they could interfere with the TATA factor interactions and prevent assembly of a stable initiation complex, and allow subsequent precise positioning of a nucleosome core on the promoter DNA between the 2 bound Adf-2 sites for stable repression.

Considerations for this model are that we have observed stable interactions of the TATA factor at the *Adh* distal promoter *in situ* in ADH⁺_A nuclei within the localized nuclease hypersensitive domain. Recent studies have suggested a role for the TATA factor binding in maintaining localized chromatin DNase I hypersensitive sites at the promoters of several *Drosophila* heat shock protein genes prior to heat shock activation (16, 36). *In vitro*, the mammalian TATA factor TFIID plays a critical role in preventing nucleosome-mediated transcription repression (37, 38). TFIID binds early and subsequently generates stable transcription initiation complex *in vitro* (39–42). Furthermore, in ADH⁺ nuclei wrapping of DNA around a positioned nucleosome core juxtaposes the 2 Adf-2 sites and bound proteins that could promote and further strengthen protein-protein interactions such that *Adh* gene repression is stably maintained despite the presence of the positive transcription factors. Other recent *in vivo* work on the yeast PHO5, and steroid inducible MMTV promoters also indicates a role for nucleosome positioning in gene regulation. In these cases gene activity appears to require removal or structural alteration of a nucleosome (50, 51, 52, 53).

However, we cannot yet explain how the mutually exclusive binding patterns are accomplished as all the sequence-specific DNA binding factors involved are detectable in both ADH⁺ and ADH⁺_A non-synchronous cell populations. This is not entirely unexpected as these factors also interact with other genes. The timing of DNA replication in each ADH cell type relative to the presence of DNA binding factors in the nucleus could be critical. Alternatively, other proteins, which do not bind DNA, could facilitate DNA binding of the TATA factor or Adf-2. An example is the immediately early protein of pseudorabies virus that stimulates the rate or stability of TFIID binding to promoter sequences in a chromatin reconstitution system *in vitro* (38). Further work on factors purification and cloning, their localization within the cells, and interactions with reconstituted chromatin templates is needed to address this problem.

ACKNOWLEDGEMENTS

We would like to gratefully acknowledge Dr. Amanda Simcox for generating and contributing *Drosophila* cell lines, Dr. Graham Thomas for teaching us the genomic footprinting technique. We thank Marty Gorovsky, Jerry Kaye, and Abe Worcel for helpful criticisms of the manuscript. This work was supported by a grant to C.B. from the National Institutes of Health (GM 34850). We would like to dedicate this work to the memory of Abraham Worcel.

REFERENCES

1. Benyajati, C., Spoerel, N., Haymerle, H., and Ashburner M. (1983) *Cell*, **33**, 125–133
2. Henikoff, S. (1983) *Nucleic Acids Res.*, **11**, 4735–4752
3. Savakis, C., Ashburner, M., and Willis, J.H. (1986) *Dev. Biol.*, **114**, 194–207
4. Lockett, T.J., and Ashburner, M. (1989) *Dev. Biol.*, **134**, 430–437
5. Heberlein, U., and Tjian, R. (1988) *Nature*, **331**, 410–415
6. Benyajati, C., Ayer, S., McKeon, J., Ewel, A., and Huang J. (1987) *Nucleic Acids Res.*, **15**, 7903–7920

7. Cartwright, I.L. (1987) *EMBO J.*, **6**, 3097–3101
8. Heberlein, U., England, B., and Tjian, R. (1985) *Cell*, **41**, 965–977
9. Corbin, V., and Maniatis, T. (1989) *Nature*, **337**, 279–282
10. Ayer, S., and Benyajati, C. *submitted*
11. Benyajati, C., and Dray, J.F. (1984) *Proc. Natl. Acad. Sci. USA*, **81**, 1701–1705
12. Benyajati, C. et al *in preparation*
13. Cartwright, I.L., and Elgin, S.C.R. (1988) In Kahl, G. (ed.), *Architecture of eukaryotic genes*, Weinheim Federal Republic of Germany: VCH Verlagsgesellschaft, p. 281–300.
14. Gross, D.S., and Garrard, W.T. (1988) *Ann. Rev. Biochem.*, **57**, 159–197
15. Church, G.M. and Gilbert, W. (1984) *Proc. Natl. Acad. Sci. USA*, **81**, 1991–1995
16. Wu, C. (1984) *Nature*, **309**, 229–234
17. Riley, D., and Weintraub, H. (1978) *Cell*, **13**, 281–293
18. Chao, M.V., Gralla, J., and Martinson, H.G. (1979) *Biochemistry*, **18**, 1068–1074
19. Prunell, A., and Kornberg, R.D. (1982) *J. Mol. Biol.*, **154**, 515–523
20. Parker, C.S., and Topol, J. (1984) *Cell*, **36**, 357–369
21. Gilmour, D.S., Dietz, T.J., and Elgin, S.C.R. (1988) *Mol. Cell Biol.*, **8**, 3204–3214
22. Klug, A., and Lutter, L.C. (1981) *Nucleic Acids Res.*, **9**, 4267–4283
23. Lutter, L.C. (1978) *J. Mol. Biol.*, **124**, 391–420
24. Lutter, L.C. (1979) *Nucleic Acids Res.*, **6**, 41–56
25. Sollner-Webb, B., Melchior, Jr., W., and Felsenfeld, G. (1978) *Cell*, **14**, 611–627
26. Hogan, M.E., Rooney, T.F., and Austin, R.H. (1987) *Nature*, **328**, 554–557
27. Liu, L.F., and Wang, J.C. (1978) *Cell*, **15**, 979–984
28. Sawadogo, M., and Roeder, R.G. (1985) *Cell*, **43**, 165–175
29. Hochschild, A., and Ptashne, M. (1986) *Cell*, **44**, 681–687
30. Nakajima, N., Horikoshi, M., and Roeder, R.G. (1988) *Mol. Cell Biol.*, **8**, 4028–4040.
31. Horikoshi, M., Hai, T., Lin, Y.S., Green, M.R., and Roeder, R.G. (1988) *Cell*, **54**, 1033–1042
32. Ramsey, N. (1986) *J. Mol. Biol.*, **189**, 179–188
33. Morse, R.H., and Simpson, R.T. (1988) *Cell*, **54**, 285–287
34. Brown, D.D. (1984) *Cell*, **37**, 359–365
35. Weintraub, H. (1985) *Cell*, **42**, 705–711
36. Thomas, G.H., and Elgin, S.C.R. (1988) *EMBO J.*, **7**, 2191–2201
37. Workman, J.L., and Roeder, R.G. (1987) *Cell*, **51**, 613–622
38. Workman, J.L., Abmayr, S.M., Cromlish, W.A., and Roeder, R.G. (1988) *Cell*, **55**, 211–219
39. Davison, B.L., Egly, J.M., Mulvihill, E.R., and Chambon, P. (1983) *Nature*, **301**, 680–686
40. Fire, A., Samuels, M., and Sharp, P.A. (1984) *J. Biol. Chem.*, **259**, 2509–2516
41. Reinberg, D., Horikoshi, M., and Roeder, R.G. (1987) *J. Biol. Chem.*, **262**, 3322–3330
42. Buratowski, S., Hahn, S., Guarente, L., and Sharp, P.A. (1989) *Cell*, **56**, 549–561
43. Shields, G., and Sang, J.H. (1977) *Drosophila Information Service*, **52**, 161
44. Simcox, A.A., Sobeih, M.M., and Shearn, A. (1985) *Som. Cell Mol. Gen.*, **11**, 63–70
45. Marchionni, M., and Gilbert, W. (1986) *Cell*, **46**, 133–141
46. Maxam, A.M., and Gilbert, W. (1980) *Meth. Enzymol.*, **65**, 499–560
47. Sollner-Webb, B., and Reeder, R.H. (1979) *Cell*, **18**, 485–499
48. Kreitman, M. (1983) *Nature*, **304**, 412–417
49. Biggin, M.D., and Tjian, R. (1989) *Trends in Genetics*, **5**, 377–383
50. Cordingley, M.G., Riegel, A.T., and Hager, G.L. (1987) *Cell*, **48**, 261–270
51. Richard-Foy, H., and Hager G.L. (1987) *EMBO J.*, **6**, 2321–2328
52. Almer, A., Rudolph, H., Hinnen, A., and Horz, W. (1986) *EMBO J.*, **5**, 2689–2696
53. Almer, A., and Horz, W. (1986) *EMBO J.*, **5**, 2681–2687