Characterization and purification of *Adh* distal promoter factor 2, Adf-2, a cell-specific and promoter-specific repressor in *Drosophila*

Cheeptip Benyajati*, Amy Ewel, Jeffrey McKeon, Meirav Chovav and Elvira Juan[†]

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

Received June 10, 1992; Revised and Accepted August 6, 1992

ABSTRACT

Chromatin footprinting in *Drosophila* tissue culture cells has detected the binding of a non-histone protein at +8 of the distal *Adh* RNA start site, on a 10-bp direct repeat motif abutting a nucleosome positioned over the inactive *Adh* distal promoter. Alternatively the active promoter is bound by a transcription initiation complex. We have characterized and purified a protein Adf-2 that binds specifically to this direct repeat motif 5'TCTCAGTGCA3', present at +8 and –202 of the distal RNA start site. DNase I footprinting, methylation interference, and UV-crosslinking analyses showed that both direct repeats interact *in vitro* with a nuclear protein of ≈ 120 kilodaltons (kDa). We purified Adf-2 through multiple rounds of sequence-specific DNA affinity chromatography. Southwestern analysis showed that the purified 120 kDa polypeptide binds the Adf-2 motif efficiently as a monomer or homomultimer. *In vivo* titrations of Adf-2 activity with the Adf-2 motif by transient co-transfection competitions in different *Drosophila* cell lines suggested that Adf-2 is a cell-specific repressor. Adf-2 has been detected ubiquitously *in vitro*, but is functional *in vivo* as a sequence-specific DNA binding protein and repressor only in the cells that have the inactive distal promoter. We discuss the possibility that an activation process is required for Adf-2 protein to bind DNA and function *in vivo*.

INTRODUCTION

Differential gene expression in eukaryotes can be regulated at the level of transcription initiation by interactions of *cis*-acting elements and sequence-specific *trans*-acting factors (1). As exemplified by the alcohol dehydrogenase (*Adh*) gene of *Drosophila melanogaster*, complex DNA-protein and protein-protein interactions may be involved. *Adh* is transcribed from 2 tandemly arranged promoters that are only ≈ 700 bp apart (2),

each has a distinct developmental and tissue-specific patterns (3,4), and each is regulated by a distinct enhancer (5,6,7). High levels of distal promoter transcription require the stage-specific activity of the *Adh* adult enhancer (AAE), located ≈ 450 bp upstream of the distal RNA start site (6,7), while proximal promoter transcription is enhanced by the *Adh* larval enhancer (ALE), located >1 kb further upstream from the AAE (5,6). In its native context the AAE stimulates only distal transcription and has no effect on the nearby proximal promoter (5,6,13). A specifically positioned nucleosome located between the AAE and distal promoter at –186 to –331 of the distal RNA start site has been implicated in facilitating interactions among proteins bound at these *cis*-regulatory elements (29).

At least 19 nuclease-protected sites were found on a 1.3 kb region containing the AAE, distal and proximal promoters using embryo nuclear extracts; 12 of the sites located highly conserved sequences between sibling species in which the *Adh* genes are functionally conserved (8). Several studies have now shown that most of these binding sites are indeed multiple *cis*-regulatory elements within the AAE (7,9,10,11), and the distal promoter (7,12,17, and this study).

The mere presence of specific *trans*-acting factors in the cells however, is not sufficient for transcriptional regulation. Adf-1 is a sequence-specific DNA binding factor that binds to a specific upstream binding site in the *Adh* distal promoter (–43 to –84 of the distal RNA start site) and activates transcription (12). The protein is present in the nucleus of somatic cells in embryos (14), most of which do not transcribe the distal promoter (4). Adf-1 is also present in tissue culture cell lines that do not transcribe the endogenous *Adh* gene (*ADH*[–]), but is functional in these cells in transient expression assays of exogenous *Adh* templates (7,9,15,16,17). Recently, we have shown by chromatin footprinting that *in vivo* the inactive distal promoter of *ADH*[–] cells is not occupied by Adf-1 or the TATA binding factor, and that the binding sites for these positive transcription factors (–43 to –84 for Adf-1; –26 to –32 for the TATA binding factor) are occupied instead by a positioned nucleosome (–40 to –185 of the distal RNA start site), and a non-histone protein bound

* To whom correspondence should be addressed

[†] Present address: Departament de Genetica, Facultat de Biologia, Universitat de Barcelona, Barcelona 08028, Spain

at a 10-bp direct repeat motif, present at +8 of the distal RNA start site (18).

As a first step towards the understanding of the mechanism by which these alternative, mutually exclusive DNA-protein interactions at the *Adh* distal promoter are established, we report here the study of a new sequence-specific DNA binding factor named Adf-2 that binds to the 10-bp direct repeat motif, which exists also at -202 of the distal RNA start site. The two 10-bp motifs thus flank the known distal promoter elements, the Adf-1 binding site and the distal TATA box. The motifs are also conserved in sequence and spacing in the *Drosophila melanogaster* sibling species, and bind proteins from embryo nuclear extracts (8). In a functional transient transfection assay using different tissue culture cell lines, Adf-2 appears to be a repressor of distal transcription, but only in ADH⁻ cells containing the inactive distal promoter, and not in ADH⁺ cells transcribing distal transcripts.

MATERIALS AND METHODS

Cells

Drosophila tissue culture cells were grown in M3 medium (for ADH⁺ cells, 1006-2), Schneider's *Drosophila* medium (for ADH⁻ cells, Schneider line 2, SL2) containing 10% fetal calf serum, and D-22 medium without serum (for ADH⁻ cells, K_{c0}) at 23°C, as described previously (16,18). ADH transcripts are not detected in ADH⁻ cells whereas cells of ADH⁺ lines transcribe their endogenous genes specifically from the distal *Adh* promoter (16).

Nuclear extracts

Nuclei were prepared from continuous cell lines according to (19), with minor modifications. Approximately 5 × 10⁹ cells (1 liter) were disrupted in a motor driven Potter-Elvehjem homogenizer (set at 9, 20 strokes) in buffer A (15 mM HEPES, pH 7.6, 60 mM KCl, 15 mM NaCl, 0.25 mM MgCl₂, 0.5 mM EGTA, 1 mM DTT, 0.5 mM spermine, 0.15 mM spermidine) with 1 mM PMSF, and 10 µg/ml each of aprotinin, leupeptin, pepstatin, TPCK, p-chloromercuriphenyl sulfonic acid. The homogenate was layered onto a 9 ml cushion of 1.7 M sucrose in buffer A (stirred interface), and centrifuged at 25,000 × g in an HB-4 rotor for 30 min at 4°C. The pelleted nuclei were washed with lysis buffer (20 mM Tris-HCl, pH 8.0, 1.5 mM MgCl₂, 1 mM DTT, 20% glycerol, 1 mM PMSF), and resuspended in the same buffer, followed by the addition of 3 M KCl to a final concentration of 0.3 M KCl. The suspension was rotated gently at 4°C for 15 min then centrifuged at 13,000 × g for 15 min. The supernatant was dialyzed against 2 × 200 ml of buffer C (25 mM HEPES, pH 7.6, 40 mM KCl, 0.1 mM EDTA, 1 mM DTT, 10% glycerol, (20)) for 6 hr at 4°C. The dialysate was cleared by centrifugation at 13,000 × g for 15 min. The supernatant was aliquoted, frozen on dry ice, and stored at -70°C. The protein concentration at this stage was usually between 10–20 mg/ml.

Purification of Adf-2

One gram of nuclear extract from K_{c0} cells was precipitated with 0.277 g of ammonium sulfate per ml. The precipitate was resuspended and dialyzed against 0.1 M HEMG (0.1 M KCl, 25 mM HEPES-KOH, pH 7.6, 0.1 mM EDTA, 6.25 mM MgCl₂, 10% glycerol, 1 mM DTT, 1 mM PMSF, and other protease inhibitors as indicated above) and centrifuged at

12,000 × g for 10 min at 4°C. The mixture was loaded on a 200 ml heparin-agarose (Affi-Gel; Bio-Rad) column. The column was washed with 3 column volumes of 0.1 M HEMG, 3 column volumes of 0.4 M HEMG (H0.4), and 3 column volumes of 1.0 M HEMG. The H0.4 fraction was pooled, dialyzed to 0.23 M HEMG, then sheared calf thymus DNA (1 µg/100 µg of protein) and Nonidet P-40 (0.1% final concentration) was added. After incubation on ice for 15 min, the solution was centrifuged for 10 min at 10,000 rpm in a SS34 rotor (Sorvall). The cleared supernatant was loaded on a 4 ml oligo-affinity column, prepared as described (21) except that CNBr-activated Sepharose (Pharmacia) was used. The annealed oligonucleotides were the 32-bp sequence of the Adf-2a site (-215 to -184, see Fig.7), 5'-gatcTCACTTATTGTTTCTCAGTGCACCTTCTGGT3', and 5'-gatcACCAGAAAGTGCAGTGCAGAAACAAATAAGTGA3'. The column was washed with 5 column volumes of 0.23 M HEMG + N (HEMG buffer with Nonidet P-40 at 0.1%), 5 column volumes of 0.4 M HEMG + N, and 5 column volumes of 1.0 M HEMG + N. The 0.4 M eluate from the 4 ml oligo-affinity column was diluted to 0.23 M HEMG + N, calf thymus DNA added, and the mixture passed over a 0.2 ml oligo-Sepharose column. The fractions were stored at -70°C.

Probes and competitors for gel electrophoresis mobility shift assays

Probes used were ³²P-radioactively end-labeled using T4 polynucleotide kinase, and subsequently digested with a second restriction enzyme (see figure legends). Unlabeled DNA competitors were obtained in a similar manner. The end-labeled probes and competitors were isolated from 5% native polyacrylamide gels by electroelution in 0.5 × TBE inside dialysis bags. Alternatively, a 36-mer synthetic double-stranded oligonucleotide, corresponding to -215 to -184 (the Adf-2a site) with the sequence GATC at the 5' ends of both strands, was also used as a specific competitor. The specific activities of end-labeled probes ranged from 5,000–15,000 cpm/fmole of DNA. Competitor DNA was quantified by comparison after electrophoresis and ethidium bromide staining to serial dilutions of DNA of known concentration digested with the same enzymes that generated the competitor fragment.

Gel mobility shift DNA-binding assay

Binding reactions (20 µl) were carried out in 32.5 mM HEPES, pH 7.6, 0.05 mM EDTA, 6.25 mM MgCl₂, 0.5 mM DTT, 50 mM KCl, 5% glycerol, and 2 µg poly(dI.dC)-poly(dI.dC) (Pharmacia), which decreases non-specific binding of protein to the labeled probes. A typical reaction contained 4–10 µg of nuclear extract and 2–6 fmole (30,000 cpm) of end-labeled DNA, which was added last after preincubating the nuclear extract for 10 min on ice. Binding proceeded for 20 min on ice. The mixture was then electrophoresed through a native 4% polyacrylamide gel (acrylamide: bis-acrylamide 29:1) containing 45 mM Tris, 45 mM boric acid, and 1 mM EDTA. Electrophoresis was carried out at 28V/cm for 3 min and at 14V/cm for 90 min to 2 hr at room temperature with buffer recirculation. For competition assays, unlabeled competitor DNA was either mixed with the probe prior to the incubation of DNA with the extract, or added 5 min after the nuclear extract and preincubated for 15 min before probe addition. In competition assays, the amount of DNA was kept constant in all the reactions with linearized BSKS+ plasmid vector (Stratagene).

DNase I footprinting analysis

For a typical preparative binding reaction, the usual conditions were scaled up 8-fold, and the amount of end-labeled DNA 8-fold, since we have observed that the yield of specific DNA-protein complex is proportional to the amount of probe used. After the binding reaction, the concentration of CaCl_2 was adjusted to 1.25 mM. DNase I (16 U/ml, Promega-Biotec) was added and the digestion was allowed to proceed for 60 sec at 25°C. Reactions were stopped by the addition of 1 μl of 0.5 M EDTA. After electrophoresis the gel was wrapped in Saran wrap and exposed wet for 1–2 hr at 4°C. The major DNA-protein complex and free DNA fragment bands were excised and electroeluted in $0.5\times\text{TBE}$. Electroelution was carried out inside dialysis bags. The electroeluate was extracted twice with phenol-chloroform, once with chloroform, once with ether, and ethanol precipitated. The pellet was rinsed thoroughly with 70% ethanol, and the products were analyzed through 8% polyacrylamide in the presence of 8M urea, followed by autoradiography at -70°C with intensifying screens.

Methylation interference assays

End-labeled DNA fragments were partially methylated at guanine residues as described in (22). Methylated DNA was ethanol precipitated twice, rinsed with 70% ethanol, and used in the binding reactions. For preparative binding, the usual conditions were scaled up 8-fold. The binding reaction, electrophoresis, and exposure were as described above. The major complex and free DNA fragments were excised and electroeluted inside dialysis bags in $0.5\times\text{TBE}$. The electroeluates were extracted twice with phenol:chloroform, once with chloroform, once with ether and ethanol precipitated, and rinsed with 70% ethanol. The dried pellet was resuspended in 100 μl of 1 M piperidine. Base cleavage reaction was carried out at 90°C followed by ethanol precipitation. The DNA fragments were analyzed by denaturing polyacrylamide gel electrophoresis as described above.

UV-crosslinking to label proteins in the gel-shift complexes

Single-stranded DNA was generated from 3 different plasmids containing the regions of interest as inserts in the multi-cloning site of pBluescriptKS+. The second DNA strand was then labeled by incorporation of BrdU and ^{32}P -dCTP using Sequenase and a sequencing primer, and following standard procedures. The appropriate DNA fragments were gel-isolated and used in the standard DNA binding reactions without or with specific and mutated specific competitors. One-half of the mixture was then electrophoresed to detect the specific DNA-protein complexes. The other half was spotted on Saran Wrap, which was placed over a 302-nm UV transilluminator, and irradiated for 20 min at 4°C. DNase I (10 U/sample), and CaCl_2 (to 10 mM) were added, and digested for 30 min at 37°C. The reaction was terminated by boiling in SDS gel sample buffer and subjected to SDS-PAGE and autoradiography. The 'Rainbow' markers (Amersham) were used as molecular size markers on the same gels.

Southwestern blotting

After SDS-polyacrylamide gel electrophoresis, proteins were electrotransferred onto a nitrocellulose filter. The filter was washed for 1–2 hr in a prebinding buffer (25 mM HEPES, pH 7.9, 6.25 mM MgCl_2 , 100 mM KCl, 1 mM DTT) containing 2.5 $\mu\text{g/ml}$ sheared calf thymus DNA, 2.5 $\mu\text{g/ml}$ denatured, sheared calf thymus DNA, and 5% nonfat dry milk at 4°C, then

probed with ^{32}P -labeled concatemerized 36-bp oligonucleotides of the Adf-2a site (10 ng/ml)(identical to the oligonucleotides used in DNA affinity column chromatography) in the same buffer containing 2.5 $\mu\text{g/ml}$ sheared calf thymus DNA, 2.5 $\mu\text{g/ml}$ denatured, sheared calf thymus DNA, and 0.25% nonfat dry milk

Figure 1. The *Drosophila Adh* gene and a partial restriction map of the distal promoter region. The striped boxes indicate the ADH protein coding exons, interrupted by two introns. The open boxes indicate the 5' and 3' transcribed, untranslated regions. D and P mark the distal (+1) and proximal RNA start sites respectively. The 10-bp direct repeat Adf-2 motifs, Adf-2a and Adf-2b, are indicated; they flank the distal promoter containing the Adf-1 binding site (–84 to –43), TATA box (–32 to –26), and distal (adult) RNA start site (+1).

Figure 2. Competition gel mobility shift assays of labeled probes from the distal promoter region. ^{32}P -labeled, BrdU-substituted DNA fragments (≈ 5 ng) were incubated with a nuclear extract made from ADH⁺, K_{c0} cells (5 μg protein per reaction). One-half of the reaction mixture was then loaded directly on a 5% polyacrylamide gel. Lanes 1–5: the probe was –248 to –130 containing the upstream 10-bp Adf-2 motif. Lanes 6–10: the probe was –21 to +53 containing the downstream 10-bp motif. Lanes 11–15: the probe was –60 to –22 containing a partial Adf-1 binding site, and the distal TATA element. Unlabeled competitor DNAs were either a synthetic double-stranded 32-bp oligomer of the intact 'Adf-2a' site, containing the region –215 to –184, or 'f-2a.Ad' a DNA fragment with the intact Adf-2a site scrambled (–201 to –215 attached to –184 to –200); the DdeI site within the Adf-2a motif was thus destroyed.

for several hr at 4°C. After 3 washes of 20 min each with the same buffer containing 0.25% nonfat dry milk and 0.1% Nonidet P-40, the nitrocellulose was dried and exposed to X-ray film.

Transient co-transfection competition in *Drosophila* cells

The calcium phosphate-DNA transfection was performed as described (15). *Drosophila* cells were seeded at $\approx 1 \times 10^7$ cells per T-25 flask in 5 ml of M3 or Schneider's *Drosophila* medium containing 10% fetal calf serum 30 hr before transfection. The reporter plasmid, pPdisADH.CAT, contains the *Drosophila Adh* distal (adult) enhancer (6,7), and distal promoter and RNA start site from the *Adh* genomic region between -661 and +51 inserted into the SmaI site of pSV₀-CAT, which lacks a known eukaryotic promoter (23). The internal control plasmid, pPacLacZ contains the *Drosophila* actin 5C promoter and RNA start site, and also the *Adh* translation initiation site fused in frame to *LacZ* (24). Competitor plasmids were 1. p(Adf-2a)₆, a multimerized 6-mer of the 32-bp Adf-2a site (between -215 to -184, with the sequence GATC at the 5' ends of both oligonucleotide strands) cloned into the BamHI site within the multi-cloning site of pBluescriptKS+ (Stratagene). Five tandem out of the six copies were in the same orientation. 2. p(Adf-2a*)₆, a multimerized 6-mer of the 29-bp mutated Adf-2a site (with a deletion of nucleotides -198, -199, -200; with the sequence GATC at the 5' ends of both oligonucleotide strands) cloned into the BamHI site within the multi-cloning site of pBSKS+. Four copies were in the same orientation. Total

supercoiled DNA transfected was 10 µg per flask; the plasmid vector pBluescriptKS+ was used as a neutral 'filler'. When multiple cell lines were being transfected in the same experiment, a single DNA-CaPO₄ precipitate mix was made, which was then distributed equally into flasks of different cell lines. Cells were harvested at 48 hr and processed for protein concentration determination, CAT enzyme activity (25) and β-galactosidase activity (24) assays.

RESULTS

Identification of a new sequence-specific DNA binding factor at the *Adh* distal promoter

We used the electrophoretic gel mobility shift assay to investigate whether any non-histone proteins extracted from nuclei prepared from ADH⁻ cells bind to DNA fragments containing the 10-bp direct repeat motifs (see Fig.1). Typically, double-stranded DNA probes were 5' end-labeled, or as shown in Fig.2, the probes were ³²P-labeled only on one DNA strand with radioactive deoxycytidine, and also with bromodeoxyuridine (BrdU) in order to photo-affinity label the bound protein subsequently. Substitution of BrdU for thymidine did not affect the sequence-specific binding activity as analyzed by gel mobility shift assays (e.g. 26,27; compare Fig.2 and Fig.4 below). Fig.2 shows that a major nucleoprotein complex formed when the DNA probes containing

Figure 3. Photo-affinity labeling of protein(s) in the gel-shift complexes. The other half of the binding reactions shown in Fig.2, was UV-irradiated, digested with DNase I, boiled in SDS gel sample buffer, and separated on a 10% SDS-polyacrylamide gel. The samples and their orders were the same as shown in Fig.2. Protein size markers were run on the same gel. A major background band of ≈ 70 –90 kDa was observed with all 3 labeled probes and with variable intensities in different experiments. The labeling of this band was not dependent on specific DNA sequences. Its origin was not determined further.

Figure 4. Competition gel mobility shift assays of labeled probes from the intact Adf-2a site, and the specifically mutated Adf-2a and Adf-2b sites. ³²P-end-labeled DNA fragments (≈ 5 ng) were incubated with a nuclear extract (5 µg of protein) made from ADH⁻, K₆₀ cells. Lanes 1–5: the probe (Adf-2a) was -248 to -130 containing the 10-bp Adf-2a motif. Lanes 6–10: the probe (Adf-2a*) was identical to the Adf-2a except with 3-bp internal deletion, from -200 to -198, and cut with different restriction enzymes in the multi-cloning site. Lanes 11–15: the probe (Adf-2b*) was -21 to +53, derived from the wildtype *Adh* distal promoter fragment containing the Adf-2b 10-bp motif, except with 3-bp internal deletion, from +10 to +12. Unlabeled competitor DNAs were as described in Fig.2.

either the upstream 10-bp motif (the a site, lane 1), or the downstream 10-bp motif (the b site, lane 6) were incubated with nuclear extract prepared from ADH⁻, K_{co} cells. These DNA-protein complexes could be competed away efficiently and equally with an unlabeled synthetic double-stranded 36-bp DNA fragment, from -215 to -184 containing the intact 10-bp motif (lanes 2-4 and lanes 7-9), while a similar-size and similar-sequence DNA fragment but with the 10-bp motif scrambled did not compete even at 20-fold molar excess (lanes 5 and 10). These results indicate that a non-histone nuclear factor present in ADH⁻ cells, which we named *Adh* distal promoter factor 2, Adf-2, binds specifically to the region -215 to -184 containing the upstream 10-bp motif. Furthermore the same factor also appears to bind specifically and with similar affinity to the fragment from site b, also containing the 10-bp motif. Consistent with the specificity of these DNA-protein interactions, different mobility shift patterns were observed (see lane 11) when the labeled DNA probe was from -60 to -22 containing previously

identified promoter elements (a partial Adf-1 binding site, -60 to -43, (12); the GAGA motif, -42 to -37, (28); and the TATA box, -32 to -26, (2)). Furthermore, the two 36-bp competitor oligomers containing either the intact or the scrambled Adf-2 sites did not compete with these gel shift patterns (lanes 12-15).

The other half of each sample in the binding reactions shown in Fig.2 was ultraviolet (UV) -irradiated to crosslink bound protein to BrdU-substituted DNA. The DNA-protein complex was digested extensively with DNase I to ensure that only short oligonucleotides from the probe DNA remained covalently crosslinked to proteins. The molecular weights of the crosslinked proteins were then determined by electrophoresis on a SDS-polyacrylamide gel. Fig.3 shows that a 110-120 kDa polypeptide was specifically ³²P-labeled by UV-crosslinking only in the binding reactions that contained the major nucleoprotein complex of Adf-2 with either the Adf-2a or the Adf-2b site (compare Fig.2 and Fig.3). This labeled polypeptide was not present when the

Figure 5. Analysis of specific Adf-2 binding at the Adf-2a site by DNase I footprinting and methylation interference. A. the sense strand of the region -248 to -130 was 5' end-labeled at the EcoRI site close to -248, and cleaved with BamHI (both were in the multi-cloning site region). B. the antisense strand of the same region was 5' end-labeled at the BamHI site close to -130, and cleaved with PstI (both in the multi-cloning site region). The labeled DNAs were used in preparative binding reactions using a nuclear extract from K_{co} cells. Lanes 1 and 5: DNase I cleavage patterns of protein-free DNA. Lanes 2 and 6: DNase I cleavage patterns of protein-bound DNA. Lanes 3 and 7: methylation patterns of protein-free DNA. Lanes 4 and 8: methylation interference patterns of protein-bound DNA. G+A: sequence markers generated by chemical cleavage (22). The brackets show the extent of DNA-protein interactions, both protection and hypersensitive sites. Closed circles indicate strong methylation interference sites while open circles indicate partial methylation interference sites.

Figure 6. Analysis of specific Adf-2 binding at the Adf-2b site by DNase I footprinting and methylation interference. A. the sense strand of the region -63 to +51 was 5' end-labeled at the BamHI site close to -63, and cleaved with PstI (both in the multi-cloning site region). B. the antisense strand of the same region was 5' end-labeled at the EcoRI site close to +51, and cleaved with BamHI (both in the multi-cloning site region). The labeled DNAs were used in preparative binding reactions using a nuclear extract from K_{co} cells. Lanes 1 and 5: DNase I cleavage patterns of protein-free DNA. Lanes 2 and 6: DNase I cleavage patterns of protein-bound DNA. Lanes 3 and 7: methylation patterns of protein-free DNA. Lanes 4 and 8: methylation interference patterns of protein-bound DNA. G+A: sequence markers generated by chemical cleavage (22). The brackets show the extent of DNA-protein interactions, both protection and hypersensitive sites. The dashed brackets mark the locations of other sequence motifs present on the same DNA strands. Closed circles indicate strong methylation interference sites while open circles indicate partial methylation interference sites.

Figure 7. Summary of DNase I footprint and methylation interference data on the Adf-2 binding sites. Underlined and overlined nucleotides are identical at the Adf-2a and the Adf-2b sites. Closed circles indicate strong methylation interference sites while open circles indicate partial methylation interference sites. Brackets indicate the nucleotides protected from DNase I cleavage. The 'symbols' indicate DNase I hypersensitive sites, with different intensity levels (filled to empty). The stippled boxes locate the nucleotides whose deletions prevent Adf-2 from binding to the remaining sequence.

DNA-Adf-2 complexes were competed away by the intact Adf-2a site, nor was it present when the -60/-22 fragment was the labeled probe (Fig.3, lanes 12-15). Thus Adf-2 may comprise at least one polypeptide of 110-120 kDa.

Similar analysis indicated that Adf-2 can be detected in nuclear extracts from several *Drosophila* cell lines, both ADH⁻ and ADH⁺. Thus, the distribution of Adf-2 appears similar to other sequence-specific DNA binding proteins of the *Adh* distal promoter; the positive transcription factors, TATA binding factor and Adf-1 were identified and partially purified from ADH⁻ cells, which do not express the endogenous *Adh* gene (20,12). However, the *in situ* analysis of DNA-protein interactions by chromatin footprinting indicate that occupation of these binding sites in nuclei is cell-type-specific and mutually exclusive, with the TATA box and Adf-1 site bound in ADH⁺ cell types and the Adf-2b site bound in ADH⁻ cell types (18).

Characterization of the *in vitro* Adf-2 binding sites

Adf-2 binding requires an intact 10-bp motif; the scrambled Adf-2a oligomer separated at the DdeI site within the 10-bp motif did not compete with the probe containing the intact site (Fig.2). We next investigated further requirements for Adf-2-DNA interactions. The Adf-2a and b sites were specifically mutated at the DdeI site within the 10-bp motif. Fig.4 shows that an internal deletion of the same 3 bp (see the sequences in Fig.7) within the 10-bp motif from both a and b sites reduced specific Adf-2-DNA complex formation drastically (lane 6, and lane 11) when compared to the intact Adf-2a site (lane 1). Competition experiments using the intact and scrambled Adf-2a oligomers as competitors identified the residual specific interactions with the remaining DNA sequences of the Adf-2 binding sites (lanes 7-10, and lanes 12-15), while not interfering with non-specific interactions.

Analyses of the Adf-2 contact sites by DNase I footprinting and methylation interference are shown in Fig.5 and Fig.6, and summarized aligned with the DNA sequence in Fig.7. For the Adf-2a site, the DNase I footprint was seen on both DNA strands

Figure 8. Scheme for purification of Adf-2 protein. The purification is described in detail in MATERIALS AND METHODS. Reproducible yields (approximate) of proteins at each step are indicated.

(Fig.5, lanes 2 and 6) over the 10-bp motif with distinct nuclease protection and hypersensitive sites. Methylation of guanine residues within the 10-bp motif on both strands prevented Adf-2 from binding (lanes 4 and 8). No further specific DNA-protein contacts were observed in the rest of the upstream region.

Specific interactions of the nucleoprotein complex containing Adf-2 and the Adf-2b site were also seen on both strands (Fig.6, lanes 2 and 6) over the 10-bp motif. Interestingly, although these nuclease protection patterns were similar to the patterns of the Adf-2a site, differences were observed even within the identical

10-bp motifs, on both strands (summarized in Fig.7), perhaps indicating some influence of the flanking DNA sequences, which share limited homology, on the overall structure of the DNA-protein complexes (see Fig.7). However, methylation interference assays, which probe the DNA-protein contacts in the major groove, showed that Adf-2 makes very similar critical contacts with the identical guanine residues within the 10-bp motif of both sites (Fig.6, lanes 4 and 8; Fig.7). Other promoter elements on the same DNA fragment (the TATA box, the GAGAGA sequence, and the partial Adf-1 binding site) show no contact with proteins under our experimental conditions.

Purification of Adf-2

The purification scheme for Adf-2 from ADH⁻, K_c0 cells is outlined in Fig.8. At every step Adf-2 sequence-specific DNA binding activity was monitored by gel mobility shift assays. The highly-enriched fractions of Adf-2 after 2 rounds of sequence-specific DNA affinity chromatography showed identical mobility shift characteristics (Fig.9A, lanes 8 and 9) to those seen with crude nuclear extracts (see Fig.2 and Fig.4), suggesting that all components required for Adf-2 binding activity have been purified. Under the rather high-salt conditions (0.23 M KCl-HEMG buffer) used in DNA affinity chromatography to reduce nonspecific binding, some Adf-2 remained in the flowthrough fraction (Fig.9A).

Three polypeptides, $\approx$ 120, 72, 43 kDa, and minor bands in between were observed consistently in the purified Adf-2 fractions (Fig.9B, lanes 8 and 9). The 120 kDa polypeptide comprised the major species, while the yields of the 72 and 43 kDa were variable. Moreover, only the 120 and 72 kDa, and minor protein bands between them, but not any bands below, bound the wildtype Adf-2a motif specifically in Southwestern analysis, and in proportion to their silver stains (Fig.9C, lanes 8 and 9), indicating they possess functional DNA binding domain(s). Taking these results together with the UV-crosslinking experiment (see Fig.3), we suggest that the isolated 120 kDa polypeptide is Adf-2 and constitutes the major nucleoprotein complex seen in gel mobility shift assays. Smaller polypeptides in the purified Adf-2 fractions are likely proteolytic products of Adf-2, and/or other tightly-associated proteins. Some, such as the 72 kDa may be bound to the Adf-2 motif and constitute other minor faster-migrating complexes seen in Fig.9A, lanes 8 and 9; the others such as the 43 kDa may be co-purified via specific protein-protein interactions with those bound to the Adf-2 motif. However, we could not eliminate the possibility that their DNA binding activity was sensitive to electrophoretic and electrotransfer conditions.

Functional activity of Adf-2

Adf-2 binding sites flank *Adh* distal promoter elements known to be involved in transcription initiation (20,12,7), therefore we next investigated the role of Adf-2 in the transcription initiation process of this promoter. Specifically, we asked whether transcription initiation of the intact *Adh* distal promoter containing the Adf-2 sites in the natural context (see Fig.1) could be affected by co-transfecting a competitor containing the Adf-2 motif. Such transient co-transfection competition assays in an ADH⁻ cell line (SL2) and an ADH⁺ cell line (1006-2) are shown in Fig.10. Clearly, expression from the distal promoter of pPdisADH.CAT in ADH⁻ cells (left panel) was specifically enhanced in *trans* by co-transfecting a plasmid carrying multimerized wildtype 32-bp Adf-2a site [p(Adf-2a)₆] in a dosage-dependent manner, but not by the plasmid carrying multimerized mutated 29-bp

Figure 9. Purified Adf-2 is a 120 kDa protein that binds the wildtype Adf-2 motif specifically. A, a portion (10 μ l/400) of every fraction across the last 0.2 ml oligo-Sepharose column was assayed by gel mobility shift using a labeled Adf-2a site as probe (see Fig. 2). Lane 1: Flowthrough. Lanes 2–7: 0.23 M KCl washes. Lanes 8–11: 0.4 M KCl eluates. Lanes 12–16: 1.0 M KCl eluates. Aff: a portion of a 0.4 M KCl pooled eluate from the 4 ml oligo-Sepharose column. F: free DNA, B: protein-bound DNA. B. and C. One-third, and two-thirds respectively of each 0.2 ml oligo-Sepharose fraction was precipitated and electrophoresed on 2 separate 4 to 15% gradient SDS-polyacrylamide gels and stained with silver (B.) or transferred to nitrocellulose and probed with ³²P-labeled concatemeric Adf-2a binding site (C.). The lanes are as indicated in panel A. The signals in C. were specific to the intact Adf-2 motif; the mutated Adf-2 motif showed much reduced binding, and several other oligonucleotides (see ref.9) did not bind at all (data not shown) M: protein size markers (approximately 100 ng/band) indicated in kilodaltons at the left in B.

Adf-2a site [p(Adf-2a*)₆]. Strikingly the enhanced expression in *trans* was observed only in ADH⁻ (left panel), and not ADH⁺ cells (right panel). Taken together these results indicate Adf-2 may play a negative role in transcription initiation from the *Adh* distal promoter *in vivo*, but in an ADH-cell-type specific manner.

DISCUSSION

In *Drosophila* tissue culture cells the *Adh* distal promoter is organized with precise positioning of non-histone proteins and nucleosome cores into either an active or inactive conformation for transcription initiation (18). The patterns of DNA-nonhistone protein interactions are mutually exclusive, with the TATA box occupied in ADH⁺ cells, and the Adf-2b site, which contains a conserved 10-bp direct repeat motif, occupied in ADH⁻ cells.

Figure 10. Transient co-transfection competition assays for Adf-2 functional activity in ADH⁻ and ADH⁺ cell lines. In each set of transfection experiments, several combinations of reporter, competitor and filler plasmid DNAs were used. For each combination, a single DNA-CaPO₄ precipitate of the reporter plasmid, pPdisADH.CAT (0.5 μ g), the internal control plasmid, pPacLacZ (0.5 μ g), the competitor plasmid(s) (0–9 μ g), and the 'filler' plasmid pBluescriptKS+ (up to 5–10 μ g total DNA) was made, and distributed equally into flasks containing (left panel) ADH⁻, SL2 cells and (right panel) ADH⁺, 1006-2 cells. Competitor plasmids were p(Adf-2a)₆, and p(Adf-2a*)₆. The CAT activity was first normalized to β -galactosidase activity to control for transfection efficiency (see ref.10), then expressed relative to the basal activity of the distal promoter construct without competition. The basal activities of this construct were 10-fold higher in ADH⁺ than ADH⁻ cells because the distal enhancer (AAE, -600 to -450) is functional only in ADH⁺ cells (7,9). The results shown represent 3 independent experiments; error bars indicate $\pm$ standard deviation.

In the present study, when the multimerized Adf-2 motif was used as a competitor by co-transfecting with the intact distal promoter linked to a CAT reporter gene, distal promoter transcription was specially enhanced in *trans* in a dosage-dependent manner, but only in ADH⁻ and not ADH⁺ cells. We suggest that in ADH⁻ cells increasing amounts of wildtype Adf-2 motif bound a repressor present in limited amount and competed it away from the 2 natural Adf-2 binding sites on the distal promoter of pPdisADH.CAT allowing transcription initiation complex formation. However, in ADH⁺ cells the Adf-2 motif had no effect over a near 10-fold range indicating the absence of the repressive binding activity. Thus based on *in vivo* observations, an active Adf-2 motif binding activity exists in ADH⁻ cells. Here we have also identified, characterized and purified from nuclear extracts of an ADH⁻ cell line a new sequence-specific DNA binding factor, Adf-2 that binds to both Adf-2 sites. Adf-2 was purified through DNA affinity chromatography using an identical Adf-2 motif that functioned in the *in vivo* competition assay. This protein is thus a good candidate for the proposed repressor.

Highly enriched Adf-2 fractions formed a single major complex on the Adf-2 motif identical in binding characteristics and gel shift mobility to those found in crude nuclear extracts. We have suggested that the purified 120 kDa polypeptide is Adf-2, and that it constitutes the major nucleoprotein complex, based on the specific photo-affinity labeling in crude nuclear extracts, its abundance in the purified fractions, and its specific binding to the Adf-2 motif in Southwestern analysis. The latter study also showed the 120 kDa polypeptide binds DNA specifically and efficiently as either a monomer or homomultimer. Interestingly, the patterns of strong DNA-protein contact sites shown in Fig.7 can be regarded as having a two-fold symmetry with the axis through nucleotides -198 (Adf-2a site) and +12 (Adf-2b) respectively. Perhaps Adf-2 is a homodimer; the 3-bp deletion

mutations, which removed the left-half protein contact sites, could still form complexes of appropriate mobility, but with the affinity reduced drastically (see Fig.4).

Adf-2 *in vitro* DNA binding activity is present equally in both ADH⁻ and ADH⁺ cells, however there was no detection of site-specific binding *in vivo* (18), nor relief of repression at the transfected distal promoter in ADH⁺ cells. The much higher basal levels of the distal promoter of pPdisADH.CAT in ADH⁺ cells (see Fig.10 legend) are also consistent with high occupancy by the transcription initiation complex at the promoter. The distal enhancer (AAE) is functional in ADH⁺ cells and stimulates transcription through the distal promoter elements, the TATA box and the Adf-1 binding site (7); low levels of transcription resulted when either or both positive distal promoter elements were deleted even in the presence of the AAE. Thus there may be an ADH⁺-cell-specific mechanism to prevent Adf-2 from binding. The suppression of Adf-2 activity *in vivo* could be due to, for example, an association with a specific inhibitory substance, or a particular protein conformation that was lost during nuclear extract preparations, or posttranslation modifications as yet undetected, but that do not affect DNA binding *in vitro*.

Based on the observed mutually exclusive protein-DNA interactions at the TATA box and the Adf-2b site *in vivo* (18), we suggested that the TATA binding protein together with transcription initiation complex, and Adf-2 (or a protein that binds at the Adf-2b site) compete for binding. Here we have shown by *in vivo* co-transfection competition that Adf-2 binding at the Adf-2 site(s) in the native context of the transfected distal promoter very likely interferes directly with transcription initiation complex formation. It is not clear, however, by which mechanism the alternative DNA-protein interactions are established in the genome and stably maintained *in vivo*. Active Adf-2 could directly prevent TFIID binding, or bound Adf-2

could position a nucleosome over the distal promoter that prevents Adf-1 and TFIID from binding. Also, our data did not distinguish whether the Adf-2a or Adf-2b sites, or both, are required for the process. Adf-2 has now been cloned (unpublished results) and functional analysis of the protein and the mechanism of interference with distal transcription initiation will be addressed.

ACKNOWLEDGMENTS

This work was supported by a grant from the National Institutes of Health (GM 34850) to C.B. Adf-2 was first discovered *in vitro* and initially characterized by Dr. Elvira Juan while she was on sabbatical at the University of Rochester. We thank Stephen Ayer and Jeffrey Jackson for critical comments on the manuscript.

REFERENCES

1. Mitchell, P.J. and Tjian, R. (1989) *Science*, 245, 371–378.
2. Benyajati, C., Spoerel, N., Haymerle, H. and Ashburner, M. (1983) *Cell*, 33, 125–133.
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. Posakony, J.W., Fischer, J.A. and Maniatis, T. (1985) *Cold Spring Harbor Symp. Quant. Biol.*, 50, 515–520.
6. Corbin, V. and Maniatis, T. (1989) *Genes Dev.*, 3, 2191–2200.
7. Ayer, S. and Benyajati, C. (1990) *Mol. Cell. Biol.*, 10, 3512–3523.
8. Moses, K., Heberlein, U. and Ashburner, M. (1990) *Mol. Cell. Biol.*, 10, 539–548.
9. Ayer, S. and Benyajati, C. (1992) *Mol. Cell. Biol.*, 12, 661–673.
10. Abel, T., Bhatt, R. and Maniatis, T. (1992) *Genes and Dev.*, 6, 466–480.
11. Falb, D. and Maniatis, T. (1992) *Genes and Dev.*, 6, 454–465.
12. Heberlein, U., England, B. and Tjian, R. (1985) *Cell*, 41, 965–977.
13. Corbin, V. and Maniatis, T. (1989) *Nature*, 337, 279–282.
14. England, B.P., Admon, A. and Tjian, R. (1992) *Proc. Natl. Acad. Sci. USA*, 89, 683–687.
15. Benyajati, C. and Dray, J.F. (1984) *Proc. Natl. Acad. Sci. USA*, 81:1701–1705.
16. Benyajati, C., Ayer, S., McKeon, J., Ewel, A. and Huang, J. (1987) *Nucleic Acids Res.*, 15, 7903–7920.
17. England, B.P., Heberlein, U. and Tjian, R. (1990) *J. Biol. Chem.*, 265, 5086–5094.
18. Ewel, A., Jackson, J.R. and Benyajati, C. (1990) *Nucleic Acids Res.*, 18, 1771–1786.
19. Varshavsky, A. (1987) *Methods in Enzymol.*, 151, 551–565.
20. Parker, C.S. and Topol, J. (1984) *Cell*, 36, 357–369.
21. Kadonaga, J.T. and Tjian, R. (1986) *Proc. Natl. Acad. Sci. USA*, 83, 5889–5893.
22. Maxam, A.M. and Gilbert, W. (1980) L. Grossman and K. Moldave (ed.) *Methods in Enzymol.*, 65, 499–560.
23. Gorman, C.M., Moffat, L.F. and Howard, B.H. (1982) *Mol. Cell. Biol.*, 2, 1044–1051.
24. Driever, W. and Nusslein-Volhard, C. (1989) *Nature*, 337, 138–142.
25. Seed, B. and Sheen, J.-Y. (1988) *Gene*, 67, 271–277.
26. Chodosh, L.A., Carthew, R.W. and Sharp, P.A. (1986) *Mol. Cell. Biol.*, 6, 4723–4733.
27. Wu, C., Walker, B., Dawid, I., Paisley, T., Zimarino, V. and Ueda, H. (1987) *Science*, 238, 1247–1253.
28. Biggin, M.D. and Tjian, R. (1988) *Cell*, 53, 699–711.
29. Jackson, J.R. and Benyajati, C. Manuscript submitted.