MOLECULAR AND CELLULAR BIOLOGY, Jan. 1990, p. 206–216
0270-7306/90/010206-11$02.00/0
Copyright © 1990, American Society for Microbiology
Vol. 10, No. 1

Regulatory Elements Mediating Transcription from the *Drosophila melanogaster* Actin 5C Proximal Promoter

YANG-TSUNG CHUNG AND ELIZABETH B. KELLER*

Section of Biochemistry, Molecular and Cell Biology, Cornell University, Ithaca, New York 14853

Received 21 August 1989/Accepted 5 October 1989

The major cytoskeletal actin gene of *Drosophila melanogaster*, the actin 5C gene, has two promoters, the proximal one of which controls constitutive synthesis of actin in all growing tissues. To locate regulatory elements required for constitutive activity of the proximal promoter, mutants of this promoter were fused to the bacterial chloramphenicol acetyltransferase gene and assayed for transient expression activity in cultured *Drosophila* embryonic Schneider line 2 cells. An essential regulatory element has been located 313 base pairs upstream from the cap site. Deletion of this element lowered expression to one-third of the wild-type level. The element has the sequence AAGTTGTAGTTG, as shown by protein-binding footprinting with the reagent methidiumpropyl-EDTA-Fe(II). This element is probably not a general one, since it was not detected in a search of the published 5′-flanking sequences of 27 *Drosophila* genes. In addition to this regulatory element, there are five GAGA elements in the actin 5C proximal promoter, some or all of which are essential for the promoter activity as shown by an in vivo competition assay. Although this promoter has no classical TATA element, there is an essential promoter region about 35 base pairs upstream from the cap site that could be a TATA surrogate. The promoter also shows sequences homologous to the alcohol dehydrogenase factor 1-binding site and to the core of the vertebrate serum response element, but mutations of these sites did not affect promoter activity in transient expression assays.

The actin 5C gene is the major gene in *Drosophila melanogaster* that encodes the cytoskeletal actin present in all cell types in all growth stages (8–10). The gene is composed of two alternative short leader exons (exons 1 and 2) and a large exon (exon 3) that contains the whole protein-coding region (2, 30). DNA transfection analyses have established that each leader exon is preceded by a separate functional promoter (3). The distal promoter controls the formation of an mRNA composed of exons 1 and 3, and the proximal promoter controls one composed of exons 2 and 3. Since the coding region lies entirely in exon 3, the actin product is the same in both cases. But the presence of two promoters provides the organism with multiple modes of regulation of this highly expressed gene (4). Whereas the proximal promoter serves the function of maintaining the normal cytoskeletal actin level in all cells, the distal promoter is developmentally regulated in different tissues and at different stages (4). With such divergent regulatory functions, it is not surprising that the sequences of these two promoters are entirely dissimilar (2, 30) except for the presence in both of two putative regulatory elements, GAGA (1, 25) and the serum response element (SRE) (3, 18, 28).

This report presents a detailed analysis of the proximal promoter and the elements required for its function. The previous study of Bond-Matthews and Davidson (3) had established that sequences responsible for much of the high activity of this promoter reside in the region between −450 and −193 relative to the cap site of exon 2 at +1 (2). In this report, we identify the responsible regulatory element by band shift, footprint, and transient expression analyses. We then show that promoter activity also depends on multiple copies of the GAGA element (1, 25) and on a region near the start of transcription. The latter may be an essential TATA surrogate in this TATA-less promoter.

* Corresponding author.

MATERIALS AND METHODS

Source and construction of plasmids. An 8.7-kilobase-pair *Eco*RI fragment of genomic DNA that contains the entire *Drosophila* actin 5C gene, pDmA2 (8), was a generous gift of E. A. Fyrberg. Expression plasmids containing fusions of the actin 5C proximal promoter to the bacterial chloramphenicol acetyltransferase (CAT) gene in the vector pUC^{PL}CAT (3) were kindly provided by B. Bond-Matthews. These plasmids have 5′ deletions at −1000, −450, and −193 and were used as the starting material for making additional 5′-deletion mutants of the promoter at −417, −335, −298, −160, and −62 fused to the CAT gene by using available restriction sites.

A plasmid with an internal deletion from −62 to −59 (designated −62/−59) was constructed by digesting the −417–CAT fusion plasmid with *Sal*I, deleting the 5′ overhang with mung bean nuclease, and self-ligating the resultant fragment. This internal deletion was checked by sequencing. Additional internal deletions of the proximal promoter, −58/−32, −109/−63, and −109/−32, were prepared by first subcloning a fragment encompassing the region to be deleted into pBluescriptSK(+) (Stratagene), making the deletion between available unique restriction sites, religating, and then reinserting the fragment into its original location in the promoter-CAT fusion plasmid. To make the probes for band shift assays, the appropriate promoter fragments were subcloned into pIBI76 and then excised by using restriction sites in the polylinker region of the vector.

Plasmid pGA-10, which contains 10 tandem repeats of a 19-base-pair (bp) fragment, 5′-TCGAGAGCGAGAGAGCG AG-3′, with the GAGA consensus (1) in a head-to-tail orientation was constructed by the method of Xiao and Lis (31). The appropriate complementary oligonucleotides (see below for source) were annealed, 5′ phosphorylated, and ligated. The resulting polymers were digested with *Xho*I and *Sal*I to eliminate all but the tandem head-to-tail repeats and then subcloned into *Xho*I-*Sal*I-digested pIBI76. The identity

of the selected product, pGA-10, was checked by sequencing. Plasmid pGA-20 with 20 copies was then made by cloning another copy of the 10-copy insert into pGA-10. Competitor B, pCP-72, has the actin 5C proximal promoter region from −160 to −62 inserted into pIBI76. Competitor C, pCD-14, has the actin 5C distal promoter region from −363 to −220 inserted into pIBI76. Plasmid pGapdh2(−937)/lacZ (26), used as an internal transfection control, was a gift of X. H. Sun.

Annealing and labeling of synthetic oligonucleotides. All oligonucleotides were synthesized by the Cornell University Biotechnology Program Oligonucleotide Synthesis Facility. To form duplexes, complementary oligonucleotides were denatured in 100 mM NaCl–10 mM Tris hydrochloride (pH 7.5)–1 mM EDTA at 90°C for 3 min and allowed to anneal by lowering the temperature to 25°C. For making a labeled probe, an annealed oligomer was end labeled with Klenow fragment plus [α-^{32}P]dATP and the other three deoxynucleoside triphosphates and purified on a 10% polyacrylamide gel. After autoradiography, the DNA band was excised from the gel and eluted overnight at 4°C in 10 mM Tris hydrochloride (pH 7.5)–1 mM EDTA–0.5 M ammonium acetate. The supernatant was run through a Sephadex G-50 spun column, and the labeled probe was precipitated with ethanol.

Preparation and fractionation of a Kc cell nuclear extract. *Drosophila* embryonic Kc cells (23) were grown in suspension culture by the Massachusetts Institute of Technology Cell Culture Center (Boston, Mass.). The crude nuclear extract was prepared as described by Parker and Topol (19). It was further fractionated on a heparin-agarose (HA) column (13) that was eluted stepwise with buffers containing 0.1, 0.25, 0.4, 0.6, and 1.0 M KCl. The 0.4 M KCl fraction, which contained the specific DNA-binding activity, was pooled and desalted by ultrafiltration on a Centricon-100 filter (Amicon Corp.). Protein concentration was determined by the Bio-Rad dye-binding assay. Crude preparations and column fractions were frozen in small samples in dry ice and stored at −70°C.

Band shift assay for DNA-binding proteins. For the band shift assay (7, 11), DNA fragments were end labeled and isolated as described above. Binding reactions were performed in a final volume of 20 μl containing the following components: 0.5 ng (5,000 to 8,000 cpm) of labeled DNA probe, 10 mM Tris hydrochloride (pH 7.6), 40 mM KCl, 1 mM EDTA, 0.5 mM dithiothreitol, and 1 to 3 μg of poly(dI-dC) · poly(dI-dC). Reactions were initiated by the addition of nuclear extract. After incubation at room temperature for 10 min, the mixture was electrophoresed through a 5% polyacrylamide gel (acrylamide-bisacrylamide, 80:1) in 6.7 mM Tris hydrochloride (pH 7.6)–3.3 mM sodium acetate–1 mM EDTA. Electrophoresis was carried out at 15 V/cm at room temperature. The gel was then transferred to Whatman 3MM paper, dried, and autoradiographed. For competition experiments, the conditions were exactly as described above except that unlabeled competitor DNA was added to the mixture before addition of the protein.

DNase I and MPE footprinting. Two-stage footprint analyses were performed as described by Carthew et al. (5), with a few modifications. The probe DNA included the genomic sequence of the actin 5C gene from −417 to −264 (Fig. 1B) that had been subcloned into pBluescriptSK(+). The probe was cut out with *Hin*dIII and *Sac*II or with *Cla*I and *Bam*HI. The two products were labeled by separate 3′-end-filling reactions, using Klenow fragment plus [α-^{32}P]dATP and the other three deoxynucleoside triphosphates. The DNA-protein binding reaction described above was scaled up fivefold, using 30 μg of protein from an HA 0.4 M KCl fraction of a Kc cell nuclear extract. After incubation at room temperature for 10 min, the binding reaction mixture was treated with either DNase I or methidiumpropyl-EDTA-Fe(II) (MPE) (15, 22).

For DNase I treatment, 1 μl of 0.5 M $MgCl_2$–0.25 M $CaCl_2$ was added to the 100-μl binding reaction mixture, followed immediately by 1 μl of 10-μg/ml DNase I. The digestion was allowed to proceed for 1 min at room temperature and terminated by bringing the EDTA concentration to 10 mM and placing the sample on ice until it could be loaded onto a 5% polyacrylamide gel (acrylamide-bisacrylamide, 30:1).

MPE treatment was carried out by adding 5 μl of freshly made 2 mM methidiumpropyl-EDTA–4 mM $Fe(NH_4)_2(SO_4)_2 \cdot 6H_2O$–80 mM dithiothreitol after binding was complete. After an additional incubation at room temperature for 10 min, the samples were immediately loaded onto a 5% gel.

Electrophoresis in both cases was performed at 15 V/cm to separate the protein-DNA complex from unbound DNA. After autoradiography of the wet gel at 4°C, the radioactive bands were excised from the gel and eluted overnight at 37°C in 0.5 M ammonium acetate–0.1% sodium dodecyl sulfate–0.1 mM EDTA–10 μg of yeast tRNA per ml at pH 7.5. The soluble fraction was extracted twice with phenol-chloroform-isoamyl alcohol (50:50:1), and the DNA was ethanol precipitated. The DNA pellet was then suspended in an 80% formamide-dye solution, denatured at 95°C for 3 min, and applied to a 6% polyacrylamide–7.5 M urea sequencing gel. Chemical cleavage of another portion of the labeled DNA at purine residues was performed as described previously (17), and samples of the product were run in parallel. After electrophoresis, the gel was transferred to Whatman 3MM paper, dried, and autoradiographed.

Cell culture, DNA transfection, and assay of CAT and β-galactosidase activities. *Drosophila* embryonic Schneider line 2 (S-2) cells (23) were grown at 23°C in Schneider medium (GIBCO Laboratories) supplemented with 10% heat-inactivated fetal bovine serum. For transfection, cells were seeded into 3 ml of fresh medium in a Falcon 25-cm^2 T flask to give a density of approximately 10^6 cells per ml. Supercoiled plasmid DNAs purified in CsCl gradients were introduced into the cells 24 h after seeding by the calcium phosphate coprecipitation method (6). Cells were left undisturbed until they were harvested. Forty-eight hours after transfection, cells were dislodged from the T flask by using a cell scraper, collected in a single Eppendorf tube by three centrifugations, washed twice with cold phosphate-buffered saline, and suspended in 200 μl of 0.25 M Tris hydrochloride (pH 7.8). The cells were disrupted by three freeze-thaw cycles of 5 min in dry ice, followed by 5 min at 37°C. Subsequently, the small pellet of cell debris was removed by centrifugation, and the supernatant was stored at −20°C.

The β-galactosidase activity of each cell extract was assayed as described by Rosenthal (20), using a 10 μl-sample of the extract and an incubation time of 15 min. The CAT activity of each cell extract was determined by the method of Gorman et al. (12), using a sample of a diluted extract equivalent to 0.5 to 1.5 μl of the cell extract and an incubation time of 10 min. All assays were performed in the linear range of CAT activity.

RESULTS

Regulatory regions in the proximal promoter of the actin 5C gene. Figure 1A shows the location of the proximal promoter upstream of exon 2 of the actin 5C gene. Transfection assays

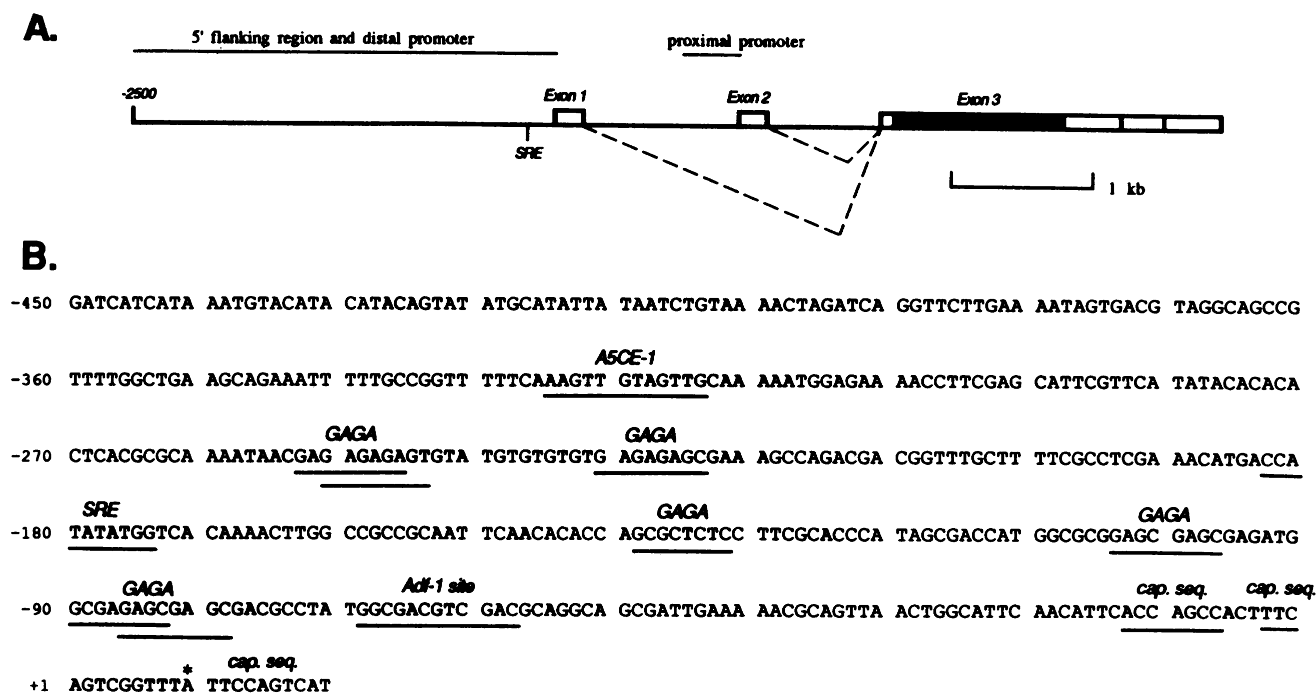

FIG. 1. Regulatory regions of the actin 5C gene. (A) Schematic diagram of the actin 5C gene. The filled box in exon 3 is the coding region. Dashed lines show how either exon 1 or exon 2 is spliced to exon 3 (2). (B) Nucleotide sequence of the actin 5C proximal promoter region and start of exon 2. Nucleotides are numbered with respect to the major cap site of exon 2 at +1 (2). The third cap site at +10 is marked (*) (2). The sequence from −22 to −350 was determined by the dideoxy-chain termination method (21). It contains two differences from the previously reported actin 5C sequences (2, 30): a G at −109 instead of C, and a C at −108 instead of G. These nucleotides were checked by restriction endonuclease cleavage at position −113 by *Nco*I and at position −106 by *Bst*UI. The rest of the sequence is taken from Bond and Davidson (2). The following putative sites for binding of transcription factors are marked: GAGA (consensus, GAGAGAGC or GCTCTCTC) (1, 25), SRE (consensus, CCATATATGG) (3, 18, 28), and Adf-1-binding site (GTCGACGTCGAC) (13). Three *Drosophila* cap consensus sequences ATCA(G/T)T(C/T) (16) and the element A5CE-1 (see text) are also marked.

had previously established that 450 bp of this proximal promoter is sufficient for full transcriptional activity (3). The sequence of this promoter is therefore given just to −450 in Fig. 1B, together with 20 bp of the start of exon 2. Several sequences homologous to the binding sites of known *Drosophila* transcription factors are marked on this sequence: five GAGA elements (1, 25), an SRE (3, 18, 28), an alcohol dehydrogenase factor 1 (Adf-1)-binding site (13), and three cap site consensus sequences (16). Also marked is the A5CE-1 element to be described in this report. There is no classical TATA element in this promoter.

The start site of exon 2 (+1 in Fig. 1B) is the primary site located by a primer extension assay by Bond and Davidson (2). A second site at +10 was also seen in the primer extension assay and in an analysis of cDNA clones (2). Such alternative sites are not unexpected in this type of TATA-less promoter. These two alternative start sites are related to the second and third cap consensus sequences (16) marked in Fig. 1B. Vigoreaux and Tobin also found two alternative start sites (30).

Binding of a Kc cell nuclear protein to the upstream region of the actin 5C proximal promoter. Bond-Matthews and Davidson reported that a 5' deletion of the promoter from −450 to −193 caused a fivefold decrease in promoter activity, demonstrating that essential regulatory elements lie in this region (3). To detect specific interactions of nuclear proteins with this upstream region, we used the band shift assay (7, 11). Three 70- to 80-bp DNA fragments spanning the sequence from −417 to −193 were end labeled as probes (Fig. 2C). Each of these probes gave retarded bands, some faint and some strong, when incubated with a crude nuclear extract from cultured *Drosophila* embryonic Kc cells in the presence of the nonspecific competitor poly(dI-dC) · poly (dI-dC) (Fig. 2A, lanes 1 to 6). DNA fragment B was chosen for further analysis because it gave the strongest shifted band (Fig. 2A, lanes 3 and 4). The major shifted band was strong even in the presence of 3 μg of poly(dI-dC) · poly (dI-dC), indicating tight specific binding of the nuclear protein in this band.

To locate the protein-binding element within probe B (−335 to −264), we carried out binding competition reactions with two components of B used as competitors: fragment D (−335 to −298) and fragment E (−298 to −264). When 25 ng of unlabeled fragment D was added to the binding reaction (Fig. 2B, lanes 9 to 11), the strong shifted band was completely lost (lane 9), whereas when fragment E was present, there was no diminution of the shifted band even at the highest concentration used (lanes 12 to 14). This result indicates that a high-affinity binding site is located within the 38-bp region between nucleotides −335 and −298.

We next fractionated the crude Kc cell nuclear extract by HA chromatography (13). A band shift assay of stepwise column fractions indicated that the protein responsible for the fragment D-binding activity eluted between 0.25 and 0.4 M KCl (the 0.4 M HA fraction). No binding activity was observed in any other fraction (data not shown).

Using the partially purified 0.4 M HA fraction, we performed a two-stage DNase I footprinting (5) of the 154-bp

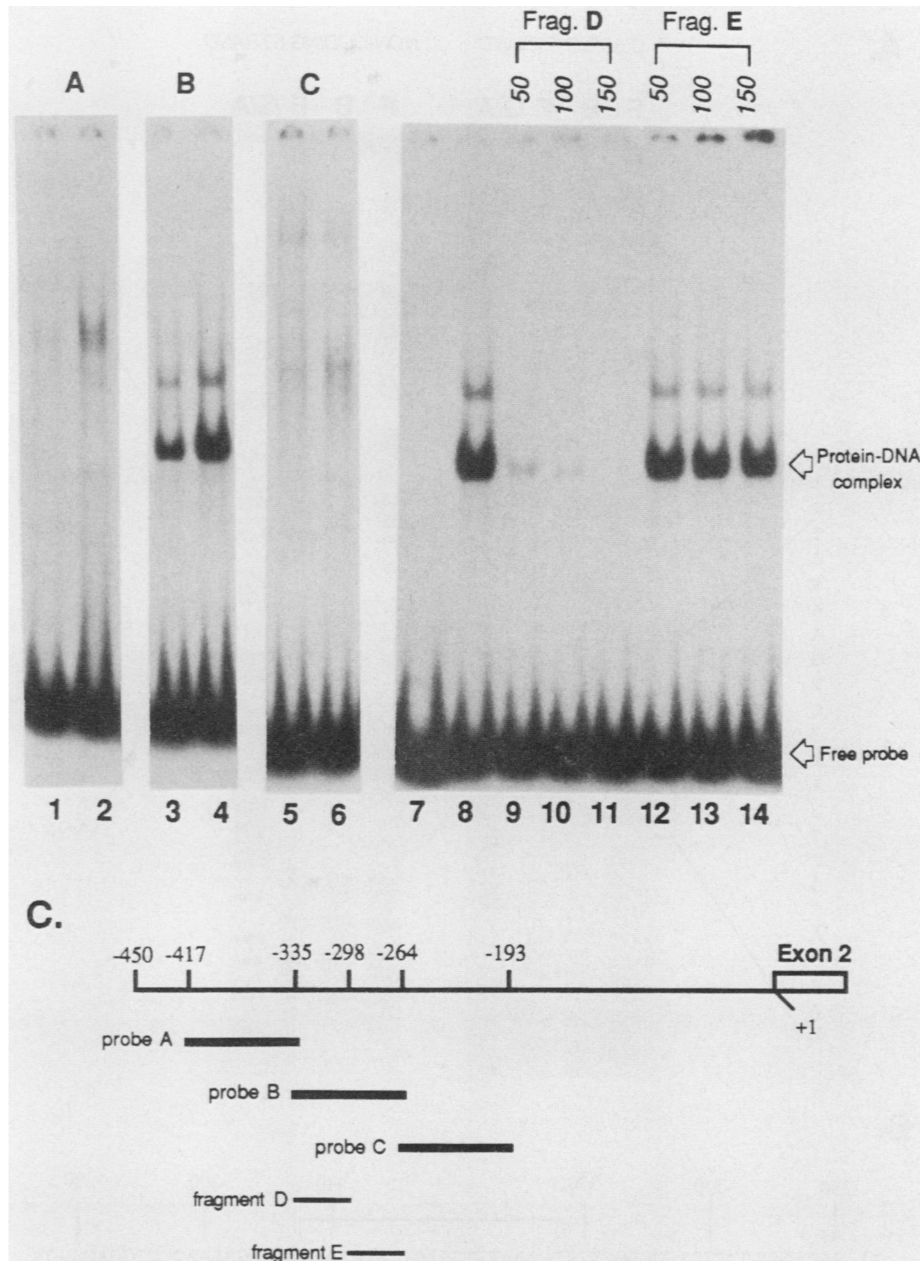

FIG. 2. Binding of a Kc cell nuclear protein to the actin 5C proximal promoter in the region between −335 and −298. The probe DNA fragments and unlabeled competitor fragments used are diagramed in panel C. All binding reaction mixtures contained 3 μg of poly(dI-dC) · poly(dI-dC) as nonspecific competitor DNA. (A) The end-labeled DNA probes (0.5 ng each) are indicated above the lanes. They were incubated with 4 μg (lanes 1, 3, and 5) or 12 μg (lanes 2, 4, and 6) of protein from a crude Kc cell nuclear extract. Binding was followed by electrophoresis through a low-ionic-strength 5% polyacrylamide gel and autoradiography. (B) Probe B was incubated without nuclear extract (lane 7) or with 5 μg of protein from a 0.4 M HA fraction of a Kc cell nuclear extract (lanes 8 to 14). Unlabeled fragment D was added to three binding reactions: 25 ng (lane 9), 50 ng (lane 10), and 75 ng (lane 11). Unlabeled fragment E was added to three others: 25 ng (lane 12), 50 ng (lane 13), and 75 ng (lane 14). The weight ratio of unlabeled fragment to labeled probe is given above lanes 9 to 14. After incubation, the reaction mixtures were electrophoresed and autoradiographed as described above.

DNA fragment encompassing the region between −417 and −264. Probes for both the coding and noncoding strands were end labeled. Protein-DNA complexes were formed, briefly digested with DNase I, and fractionated on a native polyacrylamide gel. The major shifted DNA band and the free DNA band were eluted from the gel, denatured, and electrophoresed on a standard sequencing gel. Distinct regions of complete protection were seen on each strand (Fig. 3). Protection spanned positions −328 to −308 on the coding strand and positions −332 to −312 on the noncoding strand.

To determine which nucleotides of the protected region are essential for protein binding, we did band shift assays in which mutant synthetic oligonucleotides were used as unlabeled competitors (Fig. 4A). The labeled probe WT-1 was a synthetic 33-bp oligonucleotide identical to the promoter sequence from −332 to −309. Addition of unlabeled WT-1

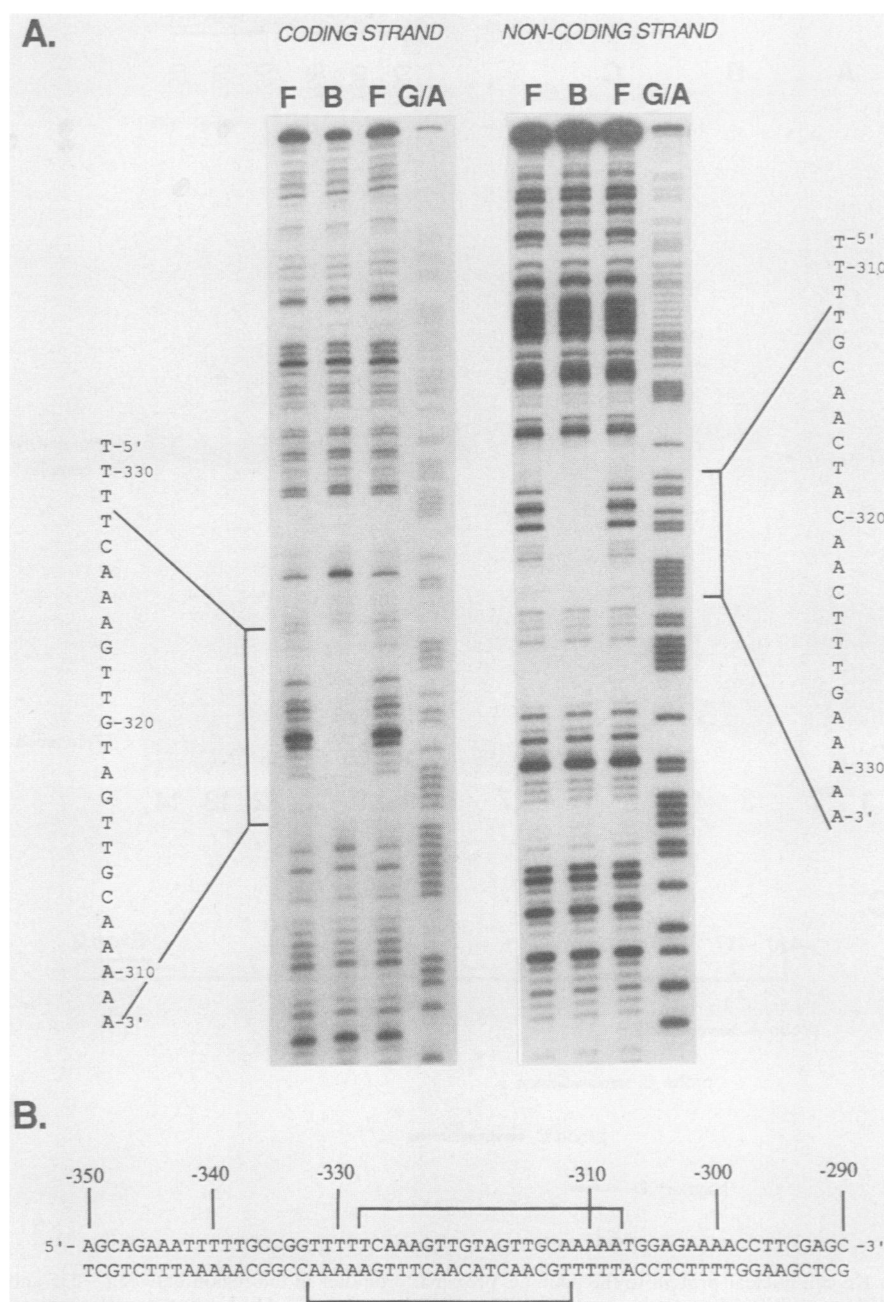

FIG. 3. DNase I footprint of a Kc cell nuclear protein bound to the actin 5C proximal promoter at −320. (A) Two-stage DNase I footprint analysis of the promoter region from −417 to −264 (see Materials and Methods). The autoradiograph of the sequencing gel shows the ladder from the free DNA band (F), from the protein-bound DNA band (B), and from a Maxam-Gilbert G+A reaction (G/A) (17) as marker. The regions protected from DNase I cleavage are bracketed. (B) Summary of the DNase I footprint analyses. The footprint regions are marked by brackets on the DNA sequence of the actin 5C proximal promoter from −290 to −350.

prevented the labeled probe from complexing with protein, as did addition of a shorter oligonucleotide identical to the promoter from −328 to −311 (WT-2) (Fig. 4B, lanes 3 and 4). Mutations near the center of this region in MT-3 and MT-4 abolished this competition (Fig. 4B, lanes 5 and 6). It appears that part or all of the sequence GTT from −323 to −321 is essential for protein binding to the element.

Finally, footprinting of the same promoter region with MPE (15, 22) was used to delineate the element more precisely. The two-stage MPE footprinting method gave a single sharply delimited region of complete protection on each strand (Fig. 5). Both footprints were 11 nucleotides long. They extended from −324 to −314 on the coding strand and from −325 to −315 on the noncoding strand. The DNA sequence protected on one or both strands is therefore 12 bp long and is AAGTTGTAGTTG. We are calling this protein-binding element the actin 5C element 1, or A5CE-1.

Generally, when a new regulatory element is identified, a number of examples are detected in different genes, and the element is then defined by a consensus. We searched the published sequences of the promoters and 5′-flanking regions of 27 *Drosophila* genes for the A5CE-1 sequence,

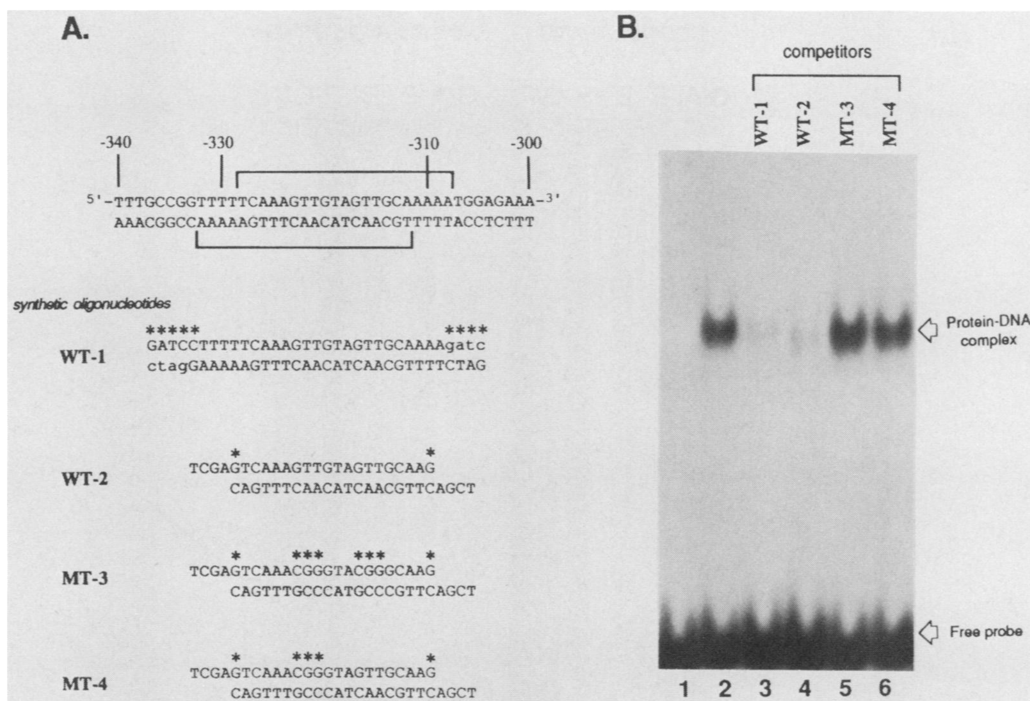

FIG. 4. Band shift assay of the element centered at −320, using wild-type and mutant elements as competitors. (A) The synthetic oligonucleotides used are aligned below the segment of the promoter from −300 to −340. The DNase I footprint of the element (Fig. 3) is marked by brackets on the promoter. Lowercase letters indicate all nucleotides incorporated into the probe WT-1 in the labeling reaction. Deviations from the wild-type sequence in duplex regions are indicated (*). (B) Band shift assays with mutant competitors. The probe was the ^{32}P-labeled synthetic wild-type fragment WT-1. It was incubated without nuclear extract (lane 1) or with 5 μg of protein from a 0.4 M HA fraction of a Kc cell nuclear extract and 1 μg of poly(dI-dC) · poly(dI-dC) (lanes 2 to 6). The following cold competitors were added to the binding reactions in an amount 100 times that of the probe: WT-1 (lane 3), WT-2 (lane 4), MT-3 (lane 5), and MT-4 (lane 6). After incubation for 10 min, reaction mixtures were fractionated through a native 5% polyacrylamide gel, and the products were visualized by autoradiography.

keeping in mind that the mutation of three nucleotides abolished protein binding to the sequence (Fig. 4). The closest match is AAGGTGTTGTTG in the promoter of an ecdysone-responsive gene *E74* (27), but it differs from A5CE-1 in two positions in the center. It is of special interest that A5CE-1 is not present in the actin 5C distal promoter (2) or 5′-flanking region extending to −2500 bp (sequence unpublished), nor is it in the essential regulatory region of the constitutive *Drosophila* glyceraldehyde 3-phosphate dehydrogenase gene 2 (*Gapdh-2*) (26). This search did not reveal any element that is identical to A5CE-1 or any that differs from it in only a single position. Until more mutational analyses are completed to determine the exact sequence requirements for A5CE-1, we have to use the MPE footprint sequence to define the element.

Transient expression analysis of actin 5C proximal promoter regulatory regions. At this point, we needed to find out whether the protein-binding element A5CE-1 is an essential regulatory element of the proximal promoter. Also, we needed to determine the contribution of the many potential general regulatory elements that can be recognized in this promoter (Fig. 1B). We adopted for this purpose the method of Bond-Matthews and Davidson (3) in which an actin 5C promoter-CAT fusion plasmid was used for transient expression assays in *Drosophila* embryonic cell cultures. A series of 5′ deletions of the actin 5C proximal promoter in this plasmid was prepared and tested in S-2 cells. To solve the problem of variations in cell numbers and transfection efficiency, we included in each transfection mixture an internal control plasmid that contains the bacterial β-galactosidase (*lacZ*) structural gene under the transcriptional control of the *Drosophila Gapdh-2* promoter (26). Also included was a carrier plasmid, pUC18, to bring the total DNA to 20 μg per flask. Forty-eight hours after transfection, cytoplasmic extracts were prepared from the transfected cells, and CAT and β-galactosidase activities were determined. The CAT activity from each transfection was corrected for differences in transfection efficiency by dividing it by the β-galactosidase activity.

In repeated in vivo assays, 5′-deletion mutants to −450, −417, and −335 all exhibited similar promoter activities that averaged 96% of the activity of the −1000 construct (data not shown). We conclude that the region distal to −335 is not important for promoter expression under these assay conditions. In the earlier band shift assay, we saw binding of a protein to the region between −417 and −335 (probe A; Fig. 2A, lanes 1 and 2), but the transfection assay results suggest that this binding does not play a role in expression under these conditions.

In contrast to the results described above, deletion of the region between −335 and −298, which has the A5CE-1 near its center, resulted in a threefold decrease in CAT activity (Fig. 6A). Since (i) the band shift assay (Fig. 2) showed that A5CE-1 is the only strong protein-binding element in this region and (ii) the DNase I footprint of A5CE-1 covered almost the entire region (Fig. 3), it is likely that A5CE-1 is the element responsible for the strong promoter activity of this region. In this case, the A5CE-1-binding protein must be a major transcription factor for constitutive expression of the actin 5C proximal promoter.

The actin 5C proximal promoter contains GAGA elements required for expression. Further deletion of the sequence

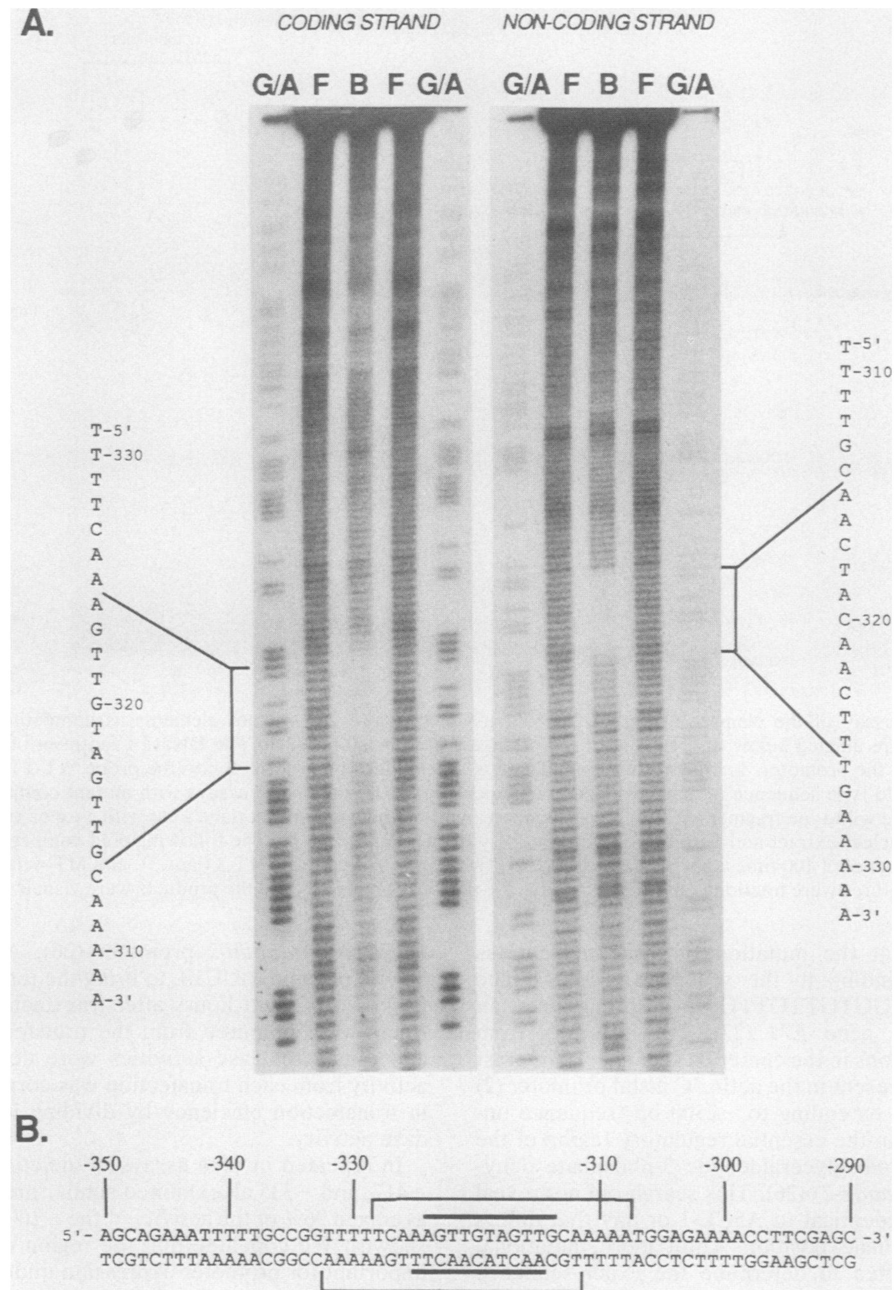

FIG. 5. MPE footprint of a Kc cell nuclear protein bound to the element centered at –320. (A) Two-stage MPE footprint analysis of the region between −417 and −264. The analysis was done as for Fig. 3 except that each protein-DNA complex was treated with an MPE mixture instead of DNase I. For each strand, the MPE ladder of the free DNA band (F) and the G+A sequencing ladder (G/A) is given on both sides of the MPE ladder from the protein-bound DNA band (B). Regions protected against MPE cleavage are bracketed. (B) Summary of the MPE footprint analyses. On the DNA sequence of the actin 5C proximal promoter from −290 to −350, the DNase I footprint (Fig. 3) is marked by brackets and the MPE footprint is marked by horizontal bars.

between nucleotides −298 and −160 decreased the promoter activity only slightly (Fig. 6A). In marked contrast, deletion of the sequence between −160 and −62 caused CAT activity to drop to 5% of full activity, suggesting that essential regulatory elements lie within this 99-bp region. Examination of this region revealed two types of sequences homologous to the binding sites of known *Drosophila* transcription factors: a potential Adf-1-binding site (13) centered at position −63 and three potential GAGA elements (1, 25) (Fig. 1B). The GAGA element was recently identified in the Ultrabithorax (*Ubx*), engrailed (*en*), and *E74* gene promoters of *D. melanogaster* (1, 25, 27). A striking common feature of these promoters is the presence of multiple repeats of the GAGA sequence motif, GAGAGAGC (1). The *Ubx* promoter possesses four copies, whereas both the *en* and the *E74* promoters possess six copies, as demonstrated by footprint analyses with the GAGA-binding protein (1, 25, 27). Similarly, the actin 5C proximal promoter contains five GAGA elements, which are centered at −86, −100, −135, −227, and −249 (Fig. 1B). All five of these elements are

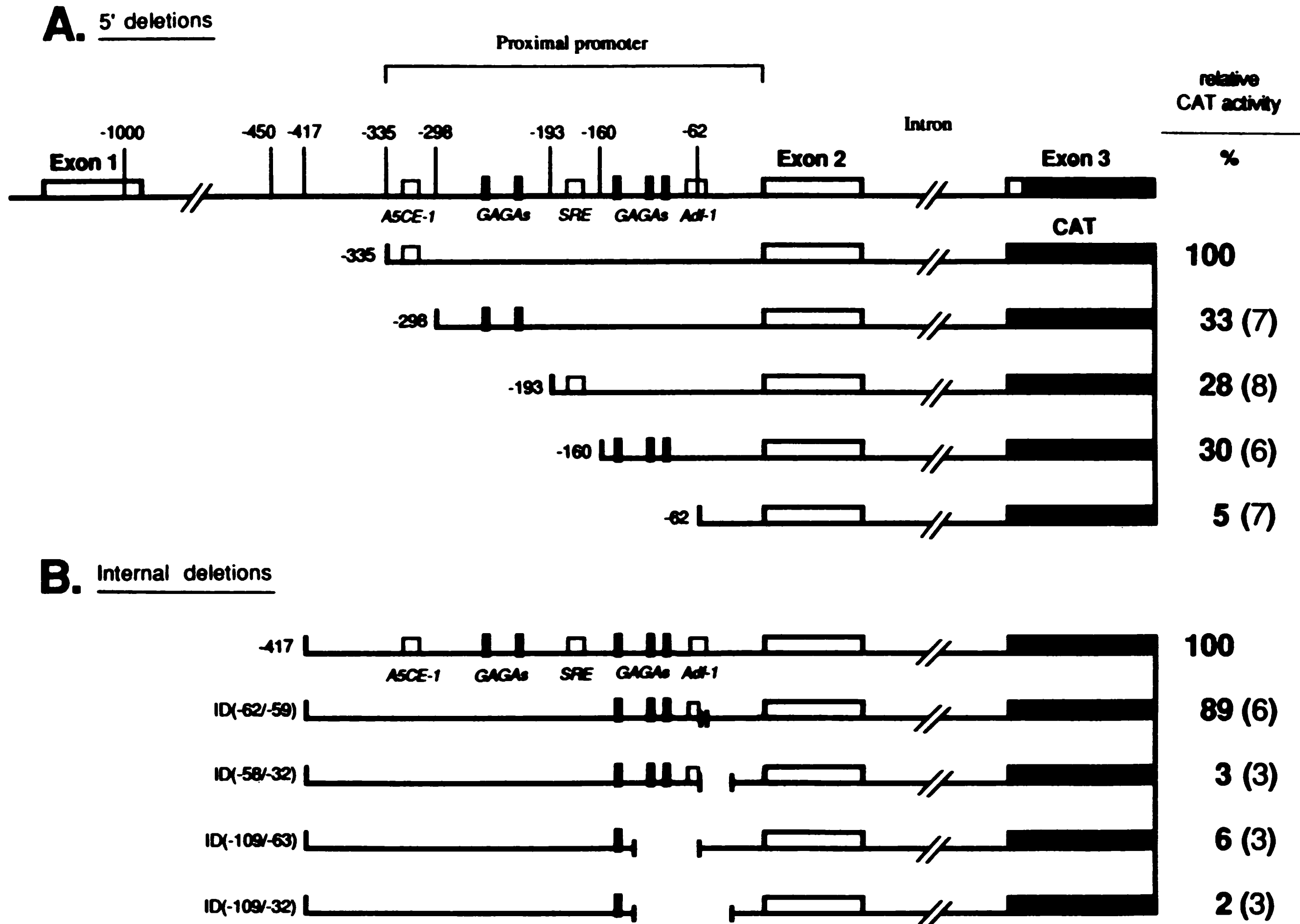

FIG. 6. Transient expression analysis of actin 5C proximal promoter mutations. (A) 5'-Deletion mutants of actin 5C proximal promoter-CAT fusion plasmids. The endpoints of the deletion mutants are indicated. The approximate locations of all potential regulatory elements are indicated schematically on the top line but are shown only at the 5' ends of the deletion mutants. Each culture of *Drosophila* S-2 cells was transfected with 1 μg of test plasmid, 2 μg of the internal control plasmid pGapdh2(−937)/lacZ (26), and 17 μg of carrier plasmid pUC18, using the calcium phosphate coprecipitation method (6). Forty-eight hours later, cellular extracts were prepared and assayed for CAT and β-galactosidase activities. The CAT activity from each flask was divided by the β-galactosidase activity to correct for differences in transfection efficiency. The 5' deletions −1000, −450, −417, and −335 all gave the same full promoter activity. The results are given as the average percentage of full promoter activity assayed in the same experiment. Each number in parentheses is the number of assays averaged. (B) Internal deletion mutants (see Materials and Methods) assayed as for panel A. All potential regulatory elements are indicated on the control plasmid, but only those elements near the deletion are shown on the mutants.

identical in sequence to authentic GAGA elements that have been footprinted by using *Drosophila* embryo cell nuclear extracts (1, 25, 27).

The drop in promoter activity on deletion of the region from −160 to −62 could be due, either entirely or in part, to the loss of GAGA elements in this region. Since specific mutagenesis of all three of the GAGAs present in this region would be technically difficult, we used an in vivo competition assay to evaluate the role of GAGAs in actin 5C transcription (24, 31). The rationale for this assay is that if an essential transcription factor is present in limited amount in the cell, it can be titrated out when cells are cotransfected with multiple copies of the binding site sequence together with the test gene. Competition for a transcription factor would depress expression of the test gene. For the assay, S-2 cells were cotransfected with various amounts of a competitor plasmid together with a constant amount of the test plasmid containing the entire actin 5C proximal promoter linked to the CAT gene. Competitor A was a plasmid with 20 tandem copies of a 19-bp synthetic fragment that includes the GAGA consensus sequence GAGAGAGC (1). This competitor caused a striking inhibition of promoter activity. With a 25-fold molar excess of the competitor A plasmid over the test plasmid, expression from the actin 5C proximal promoter was reduced to 12% of the wild-type value (Fig. 7A). It was not further reduced by a 38-fold excess. Apparently, there is a low promoter activity that is not dependent on GAGAs. Competitor B was a plasmid that contains the actin 5C proximal promoter region from −160 to −62 with three potential GAGA elements (see Fig. 1B for the sequence). This competitor also caused a marked decrease in expression of the test promoter (Fig. 7A). Competitor C was a control plasmid that contains a 144-bp fragment derived from the actin 5C distal promoter from −363 to −220 with respect to the cap site of exon 1. The sequence of this region (unpublished data) reveals no GAGA or any other recognizable element. This competitor did not influence expression of the test promoter (Fig. 7A). These results demonstrate

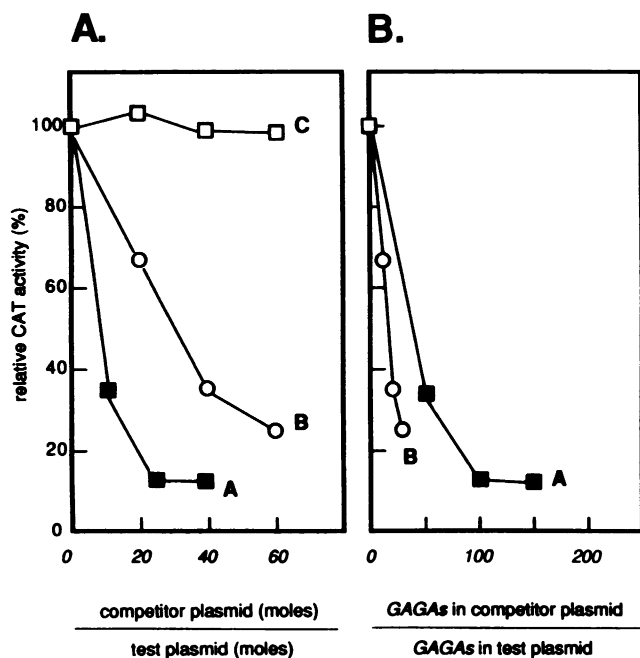

FIG. 7. In vivo competition of multiple GAGAs with the actin 5C proximal promoter. The test plasmid with the full proximal promoter-CAT fusion, −417–CAT (0.5 μg; 1 pmol), was transfected into *Drosophila* S-2 cells alone or with 12 to 60 pmol of competitor plasmids plus 1 μg of the internal control plasmid pGapdh2(−937)/lacZ (26), using the calcium phosphate coprecipitation method (6). The total amount of DNA per transfection assay was maintained at 20 μg by inclusion of the carrier plasmid pUC18. Competitor A, pGA-20, contains 20 copies of the GAGA consensus sequence (see Materials and Methods). Competitor B, pCP-72, contains one copy of the actin 5C proximal promoter region from −160 to −62 with three potential GAGAs. Competitor C, pCD-14, contains a 144-bp GAGA-free fragment of the actin 5C distal promoter from −363 to −220 relative to the cap site of exon 1. CAT and β-galactosidase activities were assayed, the CAT activity was corrected for transfection efficiency as for Fig. 6 and is given relative to the activity without competitor. It is plotted against the molar ratio of competitor to test plasmid (A) and against the ratio of GAGAs in the competitor to those in the test plasmid (B). The latter ratio is based on 20 copies of GAGA in competitor A, 3 copies in competitor B, and 5 copies in the test plasmid.

that the GAGA-binding protein plays an essential role in optimum expression of the actin 5C proximal promoter.

To compare the competition efficiency of the GAGA elements present in the synthetic poly-GAGA and the promoter fragment, we calculated the ratio of GAGAs in each of these competitors to GAGAs in the test promoter and plotted the CAT activity against this ratio (Fig. 7B). The plot showed that on the basis of GAGA content, the promoter fragment was a more efficient competitor than was the synthetic poly-GAGA. This finding could indicate that there is an additional competitive regulatory element in the region from −160 to −62 that accounts for the extra competition. It is possible, however, that the promoter GAGAs can bind the GAGA factor more effectively than can the consensus GAGA in the synthetic poly-GAGA and thus could be more competitive.

The competition results only tell us that some or all of the five GAGAs are essential. The removal of the two upstream GAGA elements located between −298 and −193 had no substantial effect on the promoter activity, as shown in the 5′-deletion assays (Fig. 6A). To assess the role of the downstream GAGA elements, we generated an internal deletion plasmid in which the promoter region from −109 to −63 was removed. The deleted region contains the two GAGA elements closest to the transcription start site (Fig. 1B). Transient expression assays showed that this operation completely eliminated the actin 5C proximal promoter activity (Fig. 6B), suggesting that one or both of these two GAGAs is essential.

The synthetic oligonucleotide with a consensus GAGA, from which the poly-GAGA was prepared, was also used for band shift assays with our Kc cell nuclear extract. Shifted bands were seen but only when we used poly(dA-dT) · poly(dA-dT) as the nonspecific competitor in place of poly(dI-dC) · poly(dI-dC) (1) (data not shown). The strong competitive inhibition of GAGA protein binding by poly(dI-dC) · poly(dI-dC) explains why we did not see significant band shifts in Fig. 2A with probe C, which includes the two upstream GAGA elements.

Role of the Adf-1-binding site in the actin 5C proximal promoter. The actin 5C proximal promoter has a sequence that differs in only one position from the prototypic Adf-1-binding site, GACGTCGACGTC, in the distal promoter of the *Drosophila* alcohol dehydrogenase gene (*adh*) (13, 14). This element is centered at −63 just as in the *adh* distal promoter (Fig. 1B). In the case of the *adh* distal promoter, when the upstream half of the palindromic Adf-1-binding site is removed by 5′ deletion, the in vitro transcriptional activity is reduced approximately twofold relative to that of the template with a full site (B. England and R. Tjian, personal communication). In our in vivo analysis of the actin 5C proximal promoter, however, when one half of the palindromic Adf-1-binding site was eliminated by deleting −62 to −59, there was no significant effect on promoter activity (Fig. 6B). It appears, therefore, that the strong activity of this actin promoter is not dependent on the Adf-1-binding site under the conditions of an in vivo assay in cultured embryonic cells.

Role of the SRE in the actin 5C proximal promoter. Both the actin 5C proximal and distal promoters have a sequence CCATATATGG (2, 4) identical to the 12-bp core of the vertebrate SRE that functions in the promoters of c-*fos*, actin, and other vertebrate genes (28). This SRE sequence could be a regulatory element in *Drosophila*, since *Drosophila* embryo cells have a nuclear protein that binds to the c-*fos* SRE, like the vertebrate serum response factor (18). A purified *Drosophila* nuclear fraction gave a footprint over the region of the actin 5C distal promoter containing the SRE in the analysis of Parker and Topol (19). Also, we found that a partially purified Kc cell nuclear protein bound to the SRE in each of the actin 5C promoters in band shift assays, and the SREs of the two promoters competed with each other for this binding (unpublished data). However, when 5′ deletions of the proximal promoter were tested in the transient expression assay, no effect was seen upon deletion of the SRE (Fig. 6A). The proximal promoter SRE apparently does not function in S-2 cells growing continuously in the presence of serum. The SREs could, however, be required for rapid initial induction of the actin 5C promoters when quiescent S-2 cells are provided with growth factors.

The actin 5C proximal promoter region near the transcription start site. An essential regulatory region lies between −58 and −32, as shown by the very low activity of the internal deletion −58/−32 plasmid (Fig. 6B). This is a region where a surrogate TATA could be located in this promoter that lacks the classical TATA box.

At the start of exon 2 are three sequences that agree with the *Drosophila* cap consensus (15). The relationship of the two start sites of Bond and Davidson (2) to two of these cap sequences is shown in Fig. 1B. These cap sequences may also be transcriptional regulatory elements.

DISCUSSION

Of the two promoters of actin 5C, the proximal promoter is the one that controls the constitutive synthesis of cytoskeletal actin in the tissues of *Drosophila* (3, 4). We have been interested in finding out how the high level of constitutive actin 5C gene expression is maintained by this promoter and what special factors make this such a powerful promoter. Previous work indicated that there is a strong regulatory element in this promoter between −450 and −193 (3). We have located within this upstream region a positive regulatory element with the sequence AAGTTGTAGTTG, which has not been previously reported and which we call the actin 5C element-1, or A5CE-1. When this element is deleted, the in vivo promoter activity drops to one-third of full activity. Thus, the A5CE-1-binding protein, which we detected in a partially purified nuclear protein fraction from cultured embryonic Kc cells, must be an essential factor in making this a strong constitutive promoter. Since A5CE-1 is located at −313 from the cap site, it could be an enhancer. Whether it has all of the properties of an enhancer remains to be determined. No enhancer activity was found in the region extending upstream from A5CE-1 to −2.5 kilobase pairs from the start of exon 1 (unpublished data).

Since the distal promoter and 5′-flanking sequences control the developmentally regulated expression of the actin 5C gene starting with exon 1 (3, 4), it is not surprising that the A5CE-1 is absent from these sequences (unpublished data). Also, we did not find any close homolog of A5CE-1 in the published 5′-flanking sequences of 26 *Drosophila* developmental genes and genes for specific tissue proteins. Therefore, the strong constitutive element A5CE-1 is probably not a general regulatory element. It is possible that it is used in a number of highly expressed constitutive genes. It is not present, however, in the promoter region controlling the high-level expression of the constitutive *Gapdh-2* gene (26).

The A5CE-1 sequence was read from a sharply defined protein-binding footprint made with the reagent MPE. The active agent from this reagent is the hydroxyl radical, so an MPE footprint locates a region of very tight binding of protein to the deoxyribose-phosphate backbone of DNA (29). In the A5CE-1 footprint, tight protein binding protects both strands over a 10-bp central region and extends 3′ one more nucleotide on each strand (Fig. 5). This is unusually long for the essential core of a nonpalindromic element. The footprint is also unusual in being an imperfect tandem repeat. The repeat of (A/T)AGTTG could bind two head-to-tail copies of a protein, but if so, both copies must be required for stable binding, since only one strong band was found in a band shift assay (Fig. 2), and it was this strong band that showed the long MPE footprint (Fig. 5).

In addition to A5CE-1, there are several types of potential general regulatory elements present in the actin 5C proximal promoter (Fig. 1B). There are five GAGA sequences that agree well with the consensus GAGAGAGC (1). Some or all of these are essential for promoter activity, as shown by an in vivo competition assay with poly-GAGA (Fig. 7A). Multiple GAGA elements have been found to function in the promoters of a number of genes of an entirely different sort, genes that are expressed only under special conditions, such as *Ubx* (1), *en* (25), and the ecdysone-responsive gene *E74* (27). In these cases, the GAGA-binding protein is not the regulator of the tissue- or stage-specific expression of the genes but an adjunct to the specific regulators (25). The same could be true of the role of the GAGA-binding protein in the regulation of actin 5C proximal promoter expression.

It is interesting that in all four of these genes in which the function of multiple GAGAs has been established, the promoter lacks a classical TATA element. But many *Drosophila* gene promoters contain sequences that agree with the GAGA consensus together with a classical TATA element. The GAGA-protein complex apparently can function with a TATA surrogate or with a TATA.

Several other elements were identified in the promoter by comparison with consensus sequences of characterized elements. One of these, the Adf-1-binding site (13), which is centered at −63 just as it is in the *adh* distal promoter, apparently does not contribute to the expression of this promoter in the transfected S-2 cells, since deletion of one half of the palindromic site had only a slight effect on expression (Fig. 6B). Another putative element is the SRE (3, 18, 28) centered at −178. This particular SRE can bind an embryonic Kc cell nuclear protein, as shown by a band shift assay (unpublished data), but our 5′-deletion assay indicated that it does not function in expression of the actin 5C proximal promoter in S-2 cells growing in the presence of serum (Fig. 6A). It is possible that this SRE, as well as the one in the distal promoter, is important for rapid induction of the actin expression when quiescent cells are induced to grow by addition of growth factors, as in the case of vertebrate actin genes (18). A role for this element only under such special circumstances could explain its presence in these two promoters, which are usually regulated differentially.

In addition to A5CE-1 and the GAGAs, there is apparently another essential element that is located just upstream from the site where a TATA box is usually found. Deletion of the sequence from −58 to −32 almost completely destroyed promoter activity in S-2 cells (Fig. 6B), and in a band shift assay, a Kc cell nuclear protein was found to bind specifically to a portion of this region between −47 and −34 (unpublished data). We did not see any band shift when we used the segment from −38 to −8 as a probe. Since Parker and Topol (19) saw a footprint of the TATA box of the actin 5C distal promoter with a similar Kc cell nuclear protein fraction, we tentatively conclude that the TTAA at −29 (2) and the nearby CATTC are not functional TATAs. The essential element between −58 and −32 may therefore be a TATA surrogate responsible for initiation of exon 2 (2, 30).

ACKNOWLEDGMENTS

We thank Beverley Bond-Matthews, Eric Fyrberg, Xiao-Hong Sun, and Bruce England for generously providing plasmids and unpublished data, Peter Dervan for a supply of methidiumpropyl-EDTA and for helpful information, Karen Kindle, Olga Perisic, and Wonkeun Lee for helpful advice, Yixin Wang, Leyla Mohassessey, and Lisa Gangarosa for help with early experiments, and John Lis for the use of cell culture facilities. We also thank the Cornell University Biotechnology Program Oligonucleotide Synthesis Facility for synthesizing oligonucleotides and the Massachusetts Institute of Technology Cell Culture Center for growing Kc cells.

This work was supported by Public Health Service grant GM10791 from the National Institutes of Health.

LITERATURE CITED

1. **Biggin, M. D., and R. Tjian.** 1988. Transcription factors that activate the *Ultrabithorax* promoter in developmentally staged

extracts. Cell **53:**699–711.

2. **Bond, B. J., and N. Davidson.** 1986. The *Drosophila melanogaster* actin 5C gene uses two transcription initiation sites and three polyadenylation sites to express multiple mRNA species. Mol. Cell. Biol. **6:**2080–2088.
3. **Bond-Matthews, B., and N. Davidson.** 1988. Transcription from each of the *Drosophila act5C* leader exons is driven by a separate functional promoter. Gene **62:**289–300.
4. **Burn, T. C., J. O. Vigoreaux, and S. L. Tobin.** 1989. Alternative *5C* actin transcripts are localized in different patterns during *Drosophila* embryogenesis. Dev. Biol. **131:**345–355.
5. **Carthew, R. W., L. A. Chodosh, and P. A. Sharp.** 1985. An RNA polymerase II transcription factor binds to an upstream element in the adenovirus major late promoter. Cell **43:**439–448.
6. **Di Nocera, P. P., and I. B. Dawid.** 1983. Transient expression of genes introduced into cultured cells of *Drosophila*. Proc. Natl. Acad. Sci. USA **80:**7095–7098.
7. **Fried, M., and D. Crothers.** 1981. Equilibria and kinetics of lac repressor-operator interactions by polyacrylamide gel electrophoresis. Nucleic Acids Res. **9:**6505–6525.
8. **Fyrberg, E. A., B. J. Bond, N. D. Hershey, K. S. Mixter, and N. Davidson.** 1981. The actin genes of *Drosophila*: protein coding regions are highly conserved but intron positions are not. Cell **24:**107–116.
9. **Fyrberg, E. A., K. L. Kindle, N. Davidson, and A. Sodja.** 1980. The actin genes of *Drosophila*: a dispersed multigene family. Cell **19:**365–378.
10. **Fyrberg, E. A., J. W. Mahaffey, B. J. Bond, and N. Davidson.** 1983. Transcripts of the six *Drosophila* actin genes accumulate in a stage and tissue specific manner. Cell **33:**115–123.
11. **Garner, M., and A. Revzin.** 1981. A gel electrophoresis method for quantifying the binding of proteins to specific DNA regions. Applications to components of the *E. coli* lactose operon regulatory system. Nucleic Acids Res. **9:**3047–3060.
12. **Gorman, C. M., L. F. Moffat, and B. H. Howard.** 1982. Recombinant genomes which express chloramphenicol acetyltransferase in mammalian cells. Mol. Cell. Biol. **2:**1044–1051.
13. **Heberlein, U., B. England, and R. Tjian.** 1985. Characterization of *Drosophila* transcription factors that activate the tandem promoters of the alcohol dehydrogenase gene. Cell **41:**965–977.
14. **Heberlein, U., and R. Tjian.** 1988. Temporal pattern of alcohol dehydrogenase gene transcription reproduced by *Drosophila* stage-specific embryonic extracts. Nature (London) **331:**410–415.
15. **Hertzberg, R. P., and P. B. Dervan.** 1984. Cleavage of DNA with (methidiumpropyl-EDTA)-iron(II): reaction conditions and product analyses. Biochemistry **23:**3934–3945.
16. **Hultmark, D., R. Klemenz, and W. J. Gehring.** 1986. Translational and transcriptional control elements in the untranslated leader of the heat-shock gene *hsp22*. Cell **44:**429–438.
17. **Maxam, A., and W. Gilbert.** 1980. Sequencing end-labeled DNA with base-specific chemical cleavages. Methods Enzymol. **65:**499–560.
18. **Mohun, T., N. Garnett, and R. Treisman.** 1987. *Xenopus* cytoskeletal actin and human c-*fos* gene promoters share a conserved protein-binding site. EMBO J. **6:**667–673.
19. **Parker, C. S., and J. Topol.** 1984. A *Drosophila* RNA polymerase II transcription factor contains a promoter-region-specific DNA-binding activity. Cell **36:**357–369.
20. **Rosenthal, N.** 1988. Identification of regulatory elements of cloned genes with functional assay. Methods Enzymol. **152:**704–720.
21. **Sanger, F., S. Nicklen, and A. R. Coulson.** 1977. DNA sequencing with chain terminating inhibitors. Proc. Natl. Acad. Sci. USA **74:**5463–5467.
22. **Sawadogo, M., and R. G. Roeder.** 1985. Interaction of a gene-specific transcription factor with the adenovirus major late promoter upstream of the TATA box region. Cell **43:**165–175.
23. **Schneider, I., and A. B. Blumenthal.** 1978. *Drosophila* cell and tissue culture, p. 266–315. *In* M. Ashburner and T. R. F. Wright (ed.), The genetics and biology of *Drosophila*, vol. 2a. Academic Press, Inc., New York.
24. **Scholer, H., and P. Gruss.** 1984. Specific interaction between enhancer-containing molecules and cellular components. Cell **36:**403–411.
25. **Soeller, W. C., S. J. Poole, and T. Kornberg.** 1988. In vitro transcription of the *Drosophila engrailed* gene. Genes Dev. **2:**68–81.
26. **Sun, X. H., J. T. Lis, and R. Wu.** 1988. The positive and negative transcriptional regulation of the *Drosophila Gapdh-2* gene. Genes Dev. **2:**743–753.
27. **Thummel, C. S.** 1989. The *Drosophila E74* promoter contains essential sequences downstream from the start side of transcription. Genes Dev. **3:**782–792.
28. **Treisman, R.** 1987. Identification and purification of a polypeptide that binds to the c-*fos* serum response element. EMBO J. **6:**2781–2784.
29. **Tullius, T. D., and B. A. Dombroski.** 1986. Hydroxyl radical "footprinting": high-resolution information about DNA-protein contacts and application to lambda repressor and Cro protein. Proc. Natl. Acad. Sci. USA **83:**5469–5473.
30. **Vigoreaux, J. O., and S. L. Tobin.** 1987. Stage-specific selection of alternative transcriptional initiation sites from the *5C* actin gene of *Drosophila melanogaster*. Genes Dev. **1:**1161–1171.
31. **Xiao, H., and J. T. Lis.** 1986. A consensus sequence polymer inhibits in vivo expression of heat shock genes. Mol. Cell. Biol. **6:**3200–3206.