MOLECULAR AND CELLULAR BIOLOGY, Nov. 1988, p. 4598–4607
0270-7306/88/114598-10$02.00/0

Vol. 8, No. 11

DNA-Binding Activities of the *Drosophila melanogaster* even-skipped Protein Are Mediated by Its Homeo Domain and Influenced by Protein Context

TIMOTHY HOEY, RAHUL WARRIOR, JOHN MANAK, AND MICHAEL LEVINE*

Department of Biological Sciences, Fairchild Center, Columbia University, New York, New York 10027

Received 10 May 1988/Accepted 1 August 1988

The homeo box gene even-skipped (*eve*) encodes a 376-amino-acid protein that binds with high affinity to sequences located near the 5′ termini of the *eve* and *en* genes. The 5′ *en* sites are A+T rich and contain copies of the 10-base-pair (bp) consensus sequence T-C-A-A-T-T-A-A-A-T. In contrast, the 5′ *eve* sites are G+C rich and contain the 9-bp sequence T-C-A-G-C-A-C-C-G. Among the five different homeo box proteins that have been tested for binding, *eve* is unique in that it shows virtually equal preference for the A+T-rich 5′ *en* binding sites and the G+C-rich 5′ *eve* sites. Most of the other proteins bind with a relatively higher affinity to the *en* sites than to the *eve* sites. In an effort to identify the regions of the *eve* protein that are responsible for its efficient binding to both classes of recognition sequences, we analyzed the DNA-binding properties of various mutant *eve* proteins. These studies suggest that the homeo domain of the *eve* protein is responsible for both binding activities. However, mutations in distant regions of the protein influenced the binding behavior of the *eve* homeo domain and caused a reduction in binding to the G+C class of recognition sites. We propose that the protein context of the homeo domain can influence its DNA-binding properties.

The differentiation of the *Drosophila* segmentation pattern depends on the activities of at least 30 different homeo box genes (reviewed in references 2, 23, and 33). Nearly every one of these genes displays a unique pattern of expression during early development, which suggests that virtually every embryonic cell contains a unique combination of homeo box gene products (1, 3, 5, 7, 18, 19, 22, 26). These different permutations of gene expression are thought to play important roles in positional information and the specification of diverse pathways of morphogenesis (10). A key question regarding the homeo box gene control of *Drosophila* development is a problem of regulation: how do each of these genes come to be expressed in the right cells at the right time? Mounting evidence suggests that selective patterns of homeo box gene expression depend on cross-regulatory interactions, whereby one homeo box gene can influence the expression of others (12, 13, 17, 24, 35, 38). Since homeo box proteins have been shown to bind to specific DNA sequences (4, 15), it is possible that such cross-regulatory interactions occur at the level of transcription. Interactions between two well-characterized segmentation genes, even-skipped (*eve*) (8, 24, 27) and engrailed (*en*) (29), provide a model for examining this problem of cross-regulation.

"Weak" (or hypomorphic) *eve* alleles result in a typical pair-rule phenotype, whereby pattern elements are deleted in alternating segments of mutant embryos (28). In contrast, "strong" (or null) alleles result in the complete loss of segment boundaries within middle-body regions (27). Previous genetic studies suggest that *eve* might play a key role in a hierarchy of interactions among segmentation genes and that the eve$^-$ phenotype results not only from the lack of *eve*$^+$ products, but also as an indirect consequence of abnormal activities of other segmentation genes. Over 20 different homeo box and segmentation gene products have been localized in *eve* mutants in an effort to identify possible target genes that might be regulated by *eve*$^+$ products (3, 12, 16, 24, 32; Frasch and Levine, in preparation). These studies suggest that the *eve* and *en* genes are either directly or indirectly regulated by *eve*$^+$ activity. It is possible that regulatory interactions between *eve* and *en* are direct since their patterns of expression are tightly coupled during normal development (21).

The eve$^-$ phenotype can be explained, at least in part, on the basis of a failure to activate *en* expression. *en*$^+$ products accumulate in the posterior compartment of each segment primordium, where they are required for the establishment and maintenance of segment boundaries (5, 7, 19). There is an absence of *en* products in the middle-body regions of *eve* embryos (12, 24). This lack of *en* expression might be a direct response to the absence of *eve*, since *en*$^+$ products are normally expressed within 10 to 20 min after the *eve* protein is first detected in a series of stripes along the anteroposterior body axis of cellularizing wild-type embryos (7, 12, 19, 24). Further support for a direct interaction between *eve* and *en* is the observation that the two genes come to be expressed within exactly the same set of cells during wild-type development (21). Similar arguments suggest that *eve* might act in *trans* to influence its own expression. Each of the known *eve* mutations disrupts the *eve* pattern during early development. At least three of these mutations have been shown to map within the protein-coding sequence, suggesting that the altered *eve* patterns result from a breakdown in a *trans*-autoregulatory interaction (Frasch and Levine, in preparation). It is possible that *eve* regulates *en* or its own expression at the level of transcription, since a full-length *eve* protein binds with high affinity to specific sites within the *eve* and *en* promoters (15).

The binding of the *eve* protein to 5′ *en* and 5′ *eve* sites involves two different classes of DNA sequences. The 5′ *en* sites are A+T rich and share a 10-base-pair (bp) consensus sequence, whereas the 5′ *eve* sites are G+C rich and contain a distinct 9-bp consensus sequence (15) (summarized in Fig. 1). *eve* is unique among five different homeo box proteins

* Corresponding author.

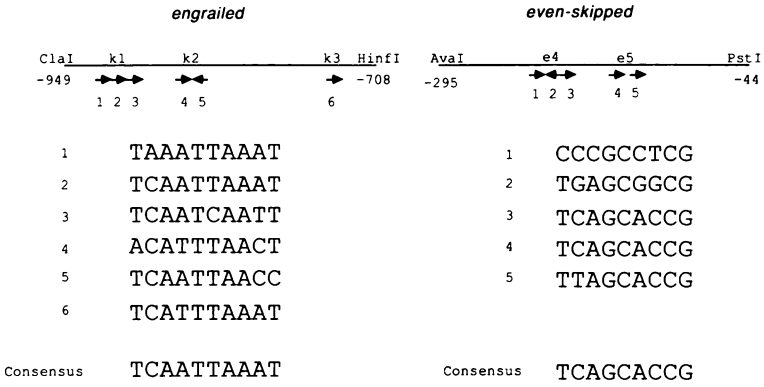

FIG. 1. *eve* protein-binding sites within the *eve* and *en* promoters. There are three binding sites clustered within a 240-bp *Cla*I-*Hin*fI fragment, located between −949 bp and −708 bp upstream from the *en* transcription start site. These sites are called k1, k2, and k3 (4, 15) and each contains at least one copy (7 of 10 matches) of the 10-bp consensus sequence T-C-A-A-T-T-A-A-A-T. The horizontal arrows indicate the locations of the consensus sequence within the k1, k2, and k3 sites. The k1 site includes three tandem copies of the consensus sequence, while the k2 site contains two copies that are oriented as inverted repeats. The nucleotide sequences of each of the indicated copies of the consensus are indicated below the map. There are two *eve* protein-binding sites, called e4 and e5, within a 250-bp *Ava*I-*Pst*I fragment located between −295 bp and −44 bp upstream from the *eve* transcription start site. Each of these sites contains multiple copies of the 9-bp consensus sequence T-C-A-G-C-A-C-C-G.

that have been examined in that it binds with about equal affinity to both classes of recognition sequences. In contrast, three of the other four proteins display a strong preference for the A+T-rich 5′ *en* sites (hereafter called the class I sites) compared with the G+C-rich 5′ *eve* sites (the class II sites). For example, the zerknullt (*zen*) (31, 37) and *en* (29) proteins each show at least a 25-fold greater preference for the class I binding sites than for the class II sites (15). The other homeo box protein that was examined, paired (*prd*) (9), contains a highly divergent homeo domain and binds with equal affinity to the class I sites and a subset of the class II sites (15).

In an effort to identify regions of the *eve* protein that are required for binding to both classes of recognition sequences, we examined the DNA-binding properties of different mutant proteins. These results suggest that the *eve* homeo domain mediates binding to both the class I and class II sites. However, the protein context of the *eve* homeo domain is important for its efficient binding to the class II sites but not to the class I sites. This result suggests that the failure of other homeo box proteins to bind to the class II sites might be due to sequence divergence both within the homeo box and outside the homeo box regions of these proteins.

MATERIALS AND METHODS

Plasmid constructions. The T7 polymerase expression vector pAR3040 (30, 36) was used to prepare each of the proteins for this study. Oligonucleotide-directed mutagenesis (39) was done to create *Nde*I restriction sites at the initiating ATGs of the *eve* and *en* coding sequences by using the oligonucleotides CGCATACCACATATGCAC and AAGTCGAAACATATGGCCCTGGAG, respectively. The pAR-*eve* expression plasmid was prepared by subcloning a 2.5-kilobase (kb) *Nde*I-*Eco*RI fragment containing the entire *eve* coding region (8) into the pAR3040 expression vector. pAR-*en* was prepared by subcloning a 1.8-kb *Nde*I-*Hin*dIII fragment containing the entire *en* coding region into pAR3040 (29). Mutant *eve* proteins were prepared by exchanging different portions of the wild-type pAR-*eve* plasmid with DNA fragments containing disruptions in the *eve* coding sequence. For the mutant $eve^{3.77.17}$ protein, the *Sac*II-*Eco*RI fragment of pAR-*eve*, which contains the coding region for amino acid residues 60 to 376 (8), was replaced with the corresponding *Sac*II-*Eco*RI fragment containing the $eve^{3.77.17}$ mutation. The $eve^{3.77.17}$ allele was obtained in a screen of a recombinant lambda library prepared with total genomic DNA from heterozygous 3.77.17 flies (R. Warrior and M. Levine, unpublished results). The *eve* Δ*Pvu*II and Δ*Nde*I proteins were prepared from pAR-*eve* derivatives that were digested with the restriction enzymes *Pvu*II and *Nde*I, respectively, and recircularized. The *eve* Bam* and Acc* proteins were prepared from pAR-*eve* derivatives that were modified by addition of a synthetic linker that contains stop codons in all reading frames after digestion with *Bam*HI or *Acc*I, respectively. The homeo box "swap" recombinant *eve* and *en* proteins were prepared from pAR-*eve* and pAR-*en* plasmids in which *Xba*I and *Sma*I sites were created on either side of the *eve* and *en* homeo boxes. The oligonucleotides GGCTCGGAGATTCTAGACGACCCGTCGGT AC and CGCCTGGCCCGGGGCAGCCGTCTAC were used to create the pAR-*eve* derivative called pAR-eve^{XS}, and the oligonucleotides GCCCCAAACAGCCTCTAGACA AGACCAACG and GGCTCCAAAAATCCCGGGGCACT GCAG were used to create the pAR-*en* derivative called pAR-en^{XS}. The *Xba*I-*Sma*I fragments containing the *eve* and *en* homeo boxes in these derivatives were exchanged, creating the recombinant pAR-eve^{ENHB} and pAR-en^{EVEHB} plasmids.

Preparation of protein extracts. Homeo box proteins were expressed in the bacterial strain BL21(DE3), which contains a single copy of the T7 RNA polymerase gene under the control of the inducible *lacUV5* promoter (36). Cells transformed with pAR3040-homeo box gene recombinant plasmids were grown at 37°C in 2× YT medium with 200 μg of ampicillin per ml and 0.4% glucose. The cells were grown to an A_{600} of 0.5, IPTG (isopropyl-β-D-thiogalactopyranoside) was added to a final concentration of 0.4 mM, and then the cells were grown for an additional 2 h.

The following steps were done on ice or at 4°C. Induced cells were harvested by centrifugation and resuspended in 1/200 volume of buffer Z (100 mM KCl, 25 mM HEPES [*N*-2-hydroxyethylpiperazine-*N*′-2-ethanesulfonic acid, pH 7.8], 12.5 mM $MgCl_2$, 1 mM dithiothreitol [DTT], 0.1%

Nonidet P-40 [NP-40; Sigma], 20% glycerol) with protease inhibitors (1 mM phenylmethylsulfonyl fluoride [PMSF], 2 mM benzamidine, leupeptin [5 μg/ml], pepstatin A [5 μg/ml], and lysozyme [0.5 mg/ml]). Cells were incubated on ice for 15 min and then lysed by sonicating twice for 15 s each at a setting 2 of a Branson sonifier. The lysate was centrifuged at 15 krpm for 10 min. The supernatant was recovered and stored at −70°C. The pellet was resuspended in buffer Z with 4 M guanidine hydrochloride and incubated for 30 min. After solubilization of the pellet, the denaturant was gradually removed by dialysis against buffer Z with 1 M guanidine hydrochloride, followed by dialysis against buffer Z alone. Insoluble material was removed by centrifugation at 15 krpm for 2 to 5 min. The DNA-binding experiments were done with the guanidine hydrochloride-extracted protein. For several of the proteins, binding was also done with the original supernatant obtained directly from induced cells. Similar results were obtained when both preparations were compared (data not shown).

Footprint assays. DNase I protection assays were done essentially as described by Heberlein et al. (14). Binding reactions were done with 2 to 5 ng of ^{32}P-labeled DNA and 5 μg of sonicated calf thymus DNA in 50 μl for 30 to 45 min on ice. The binding buffer contained 110 mM KCl, 47.5 mM HEPES, pH 7.8, 13.75 mM $MgCl_2$, 0.05 mM EDTA, 1 mM DTT, 17% glycerol, and 0.05% NP-40. After binding, 50 μl of 10 mM $MgCl_2$–5 mM $CaCl_2$ was added to the reaction, followed by 5 μl of freshly diluted DNase I (Worthington) at a final concentration of 10 μg/ml. DNase I digestion was done for 5 min on ice. The reaction was stopped by addition of 90 μl of 1% sodium dodecyl sulfate (SDS)–20 mM EDTA–200 mM KCl–250 μg of yeast tRNA per ml. The samples were extracted twice with phenol-chloroform (1:1), ethanol precipitated, and electrophoresed in 6 to 10% polyacrylamide–7.5 M urea gels (25).

DNA fragments used for footprint assays were prepared by labeling with [γ-^{32}P]ATP and polynucleotide kinase or with [α-^{32}P]dCTP and Klenow polymerase. Labeled fragments were purified after fractionation in 5% polyacrylamide gels. Portions of the labeled DNA were used for binding experiments or for sequencing by the chemical cleavage method (25).

RESULTS

The DNA-binding activities of a full-length *eve* protein have been described previously and are summarized in Fig. 1 (15). The *en* promoter contains a cluster of binding sites located between −949 bp and −708 bp upstream from the transcription start site. There are several binding sites within the proximal region of the *eve* promoter, between −295 bp and −44 bp upstream from the start of transcription. Each of the 5′ *en* sites is A+T rich and contains at least one copy of the consensus sequence T-C-A-A-T-T-A-A-A-T. In contrast, the 5′ *eve* sites are G+C rich and contain the 9-bp sequence T-C-A-G-C-A-C-C-G.

The 3.77.17 mutation uncouples binding to the class I and class II sites. In an effort to identify regions of the *eve* protein that mediate binding to the two different classes of recognition sequences, we examined the binding activities of the mutant *eve* proteins shown in Fig. 2a. Each of these proteins was prepared in a T7 expression system (36), and enriched extracts were used for DNA-binding studies (Fig. 2b). Mutant proteins were tested for binding to class I and class II sites by DNase I protection. The class I sites are located on a 240-bp *Cla*I-*Hin*fI fragment derived from the 5′ end of the *en* gene, and the class II sites are present on a 250-bp *Ava*I-*Pst*I fragment from a proximal region of the *eve* promoter (Fig. 1). These two fragments were end labeled with ^{32}P and separately incubated with different concentrations of wild-type and mutant *eve* proteins.

A small in-frame deletion within the homeo domain, called ΔPvuII (Fig. 2), completely abolished binding to both classes of recognition sequences (Fig. 3). The deletion removed 25 amino acid residues from a central region of the *eve* homeo domain, including the first 6 of the 8 residues that constitute the putative helix 2; the putative recognition helix is left intact (20). Note that a concentration of the wild-type protein that protected the 5′ *en* k2 binding site gave about equal protection of the 5′ *eve* e4 site (compare lanes 3 of Fig. 3a and b). A fivefold-higher concentration of the mutant protein failed to give detectable binding to either of these sites or to any of the other sites located within the 5′ *en* and 5′ *eve* DNA fragments. This result suggests that the *eve* homeo domain is responsible for both binding activities. However, it is also possible that the deletion within the homeo domain changes the overall configuration of the protein or in some other nonspecific manner inactivates the protein.

The 3.77.17 protein is encoded by an X-ray-induced mutant *eve* allele (27) that results from the deletion of a 24-bp region and the insertion of two extraneous nucleotides (summarized in Fig. 2). The mutation causes a frame-shift, such that the carboxyl-terminal 134 amino acid residues are replaced by 79 foreign amino acid residues. The mutant protein displayed essentially normal binding to the class I sites, in that increasing concentrations of the extract filled each of the three 5′ *en*-binding sites (lanes 6 to 8 of Fig. 4a). There was almost no difference in the footprints observed with high concentrations of the wild-type and mutant proteins (compare lanes 4 and 8, Fig. 4a). In this and subsequent experiments, we will consider only the binding of a given protein to 5′ *en* versus 5′ *eve* sites with binding to the 5′ *en* k1 and k2 sites as internal standards.

Concentrations of the wild-type *eve* protein that filled the k1 and k2 sites gave comparable protection of the 5′ *eve* e4 site, indicating that the wild-type protein possesses about an equal affinity for these sites (compare lanes 3 and 4 of Fig. 4a with lanes 3 and 4 of Fig. 4b). However, a concentration of the mutant 3.77.17 protein that was sufficient to fully protect k1 and k2 gave only partial binding to the e4 site (compare lane 8 of Fig. 4a with lane 8 of 4b). The e4 site actually consists of two half-sites, which show slightly different affinities for the wild-type *eve* protein. Note that the strong half-site within e4 was protected at lower concentrations of the wild-type protein than was the weak half-site (lane 3, Fig. 4b). The highest concentration of the mutant 3.77.17 protein that was assayed failed to bind to the weak half-site and showed at least a twofold relative reduction in binding to the strong half-site with the k1 site used as an internal standard.

Truncated proteins show normal binding to the class II sites. It is possible that the relative reduction in binding of the 3.77.17 protein to the class II sites results from the loss of the carboxyl terminus of the wild-type *eve* protein. Alternatively, perhaps this reduction results from the replacement of the carboxyl terminus with extraneous amino acid residues (Fig. 2). To distinguish between these possibilities, we tested the DNA-binding activities of two different truncated *eve* proteins that delete the carboxyl terminus. The region that has been removed from one of these mutant proteins, called Bam*, corresponds almost exactly to the region that is missing in the 3.77.17 protein. The native amino acid sequence of the 3.77.17 protein ends with residue 242, while

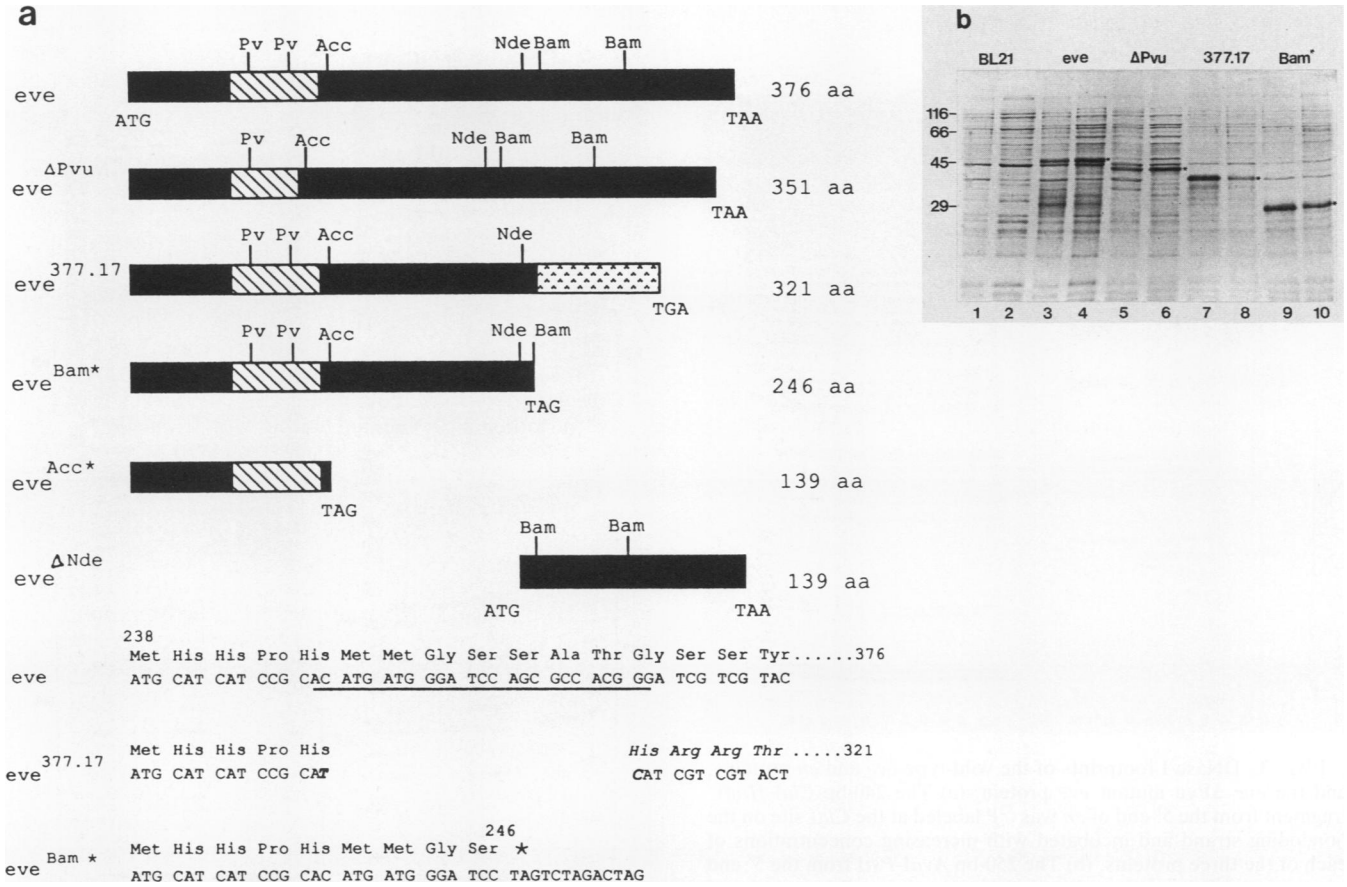

FIG. 2. Mutant even-skipped proteins. (a) Summary of the proteins examined in this study. The shaded rectangles indicate the coding sequences for the wild-type *eve* protein and various mutant *eve* proteins. The cross-hatched regions indicate the locations of the homeo box, which includes amino acid (aa) residues 70 through 130 of the wild-type protein. The ΔPvu mutant contains a small in-frame deletion within the *eve* homeo domain, which lacks residues 82 through 106 of the wild-type protein. The 3.77.17 mutation contains a 24-bp deletion and a 2-bp insertion (indicated by the italics) within 3′ coding sequences, resulting in a frame-shift of the carboxyl-terminal residues. The details of the mutation are shown at the bottom of the figure. The region that is deleted in the mutant is underlined in the wild-type *eve* sequence. This causes a frame-shift of amino acid residues 243 through 321 and a prematurely terminated protein. The extraneous residues present in the 3.77.17 protein are indicated by the stippled region in the rectangular representation of the protein shown near the top of the figure. The truncated *eve* proteins Bam* and AccI* were prepared by inserting stop codons at the *Bam* and *Acc*I sites, respectively. The Bam* protein is terminated at residue 246 of the wild-type *eve* protein, which nearly corresponds to the last native amino acid residue present in the 3.77.17 protein. The AccI* protein includes residue 139 of the wild-type *eve* protein, as well as four extraneous amino acid residues from the inserted linker. The native sequence ends just nine amino acid residues downstream from the homeo domain. The ΔNdeI protein is a carboxyl-terminal peptide of the wild-type *eve* protein, and the initiating ATG corresponds to amino acid residue 238. (b) Polyacrylamide-SDS gel of mutant proteins. Protein extracts were electrophoresed in a 12% polyacrylamide gel and stained with Coomassie blue. The odd-numbered lanes contain total protein from bacterial cells that were induced to express the T7 plasmid with IPTG (see Materials and Methods). The even-numbered lanes contain the guanidine-solubilized extracts that were used for the binding studies. Lanes 1 and 2, Control extracts containing total protein and soluble protein, respectively, prepared from BL21(DE3) cells, which contain the T7 expression vector without *Drosophila* inserts. Lanes 3 and 4, Extracts from cells expressing the full-length, wild-type *eve* protein; total extract and soluble extract used for binding studies, respectively. Lanes 5 and 6, Extracts of the ΔPvuII mutant protein. Lanes 7 and 8, Extracts of the $eve^{3.77.17}$ protein. Lanes 9 and 10, Extracts of the Bam* truncated *eve* protein. The asterisks indicate the locations of the full-length versions of each protein. Note that in every case the most prominent band corresponds to the *eve* protein of expected size. The positions of molecular mass markers (in kilodaltons) are indicated to the left of the gel.

the Bam* protein ends at residue 246 (Fig. 2). Figure 5 compares the binding of the Bam* protein with that of the wild-type *eve* protein. A concentration of the mutant protein that protected the 5′ *en* k1 and k2 sites also fully protected both the weak and strong half-sites within the 5′ *eve* e4 site. Thus, the Bam* protein displayed an equal preference for the class I and class II binding sites, similar to that observed for the wild-type *eve* protein. This observation suggests that the carboxyl terminus of the wild-type *eve* protein does not play a specific role in its binding to class I versus class II sites. Consistent with this conclusion is the observation that a second truncated *eve* protein (called AccI*; see Fig. 2), which included only the amino-terminal 139 amino acid residues of the native protein and terminated just downstream from the *eve* homeo domain, also bound equally well to the class I and class II binding sites. Moreover, a carboxyl-terminal polypeptide (called ΔNdeI; see Fig. 2) that included residues 238 through 376 of the native *eve* protein did not possess DNA-binding activity in our assays (data not shown). These results, along with the binding study shown in

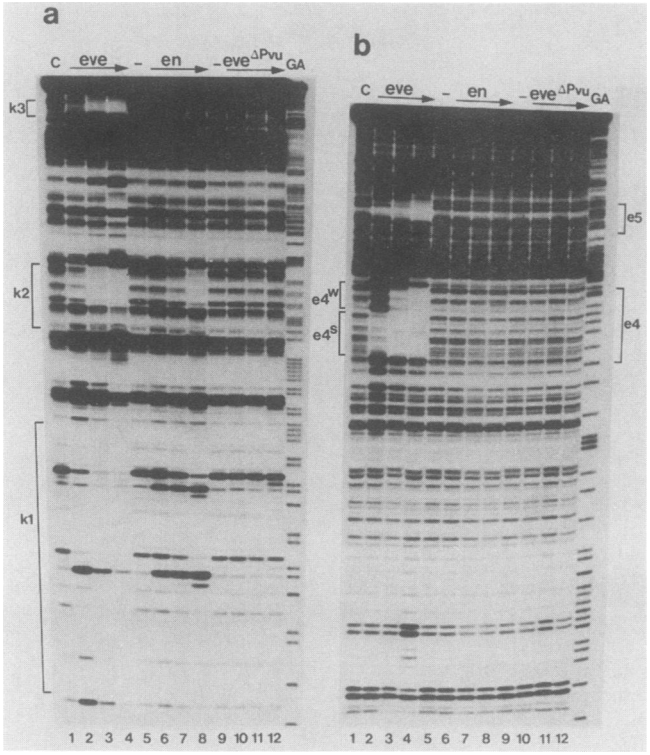

FIG. 3. DNase I footprints of the wild-type *eve* and *en* proteins, and the *eve* ΔPvu mutant *eve* protein. (a) The 240-bp *Cla*I-*Hin*fI fragment from the 5′ end of *en* was ^{32}P labeled at the *Cla*I site on the noncoding strand and incubated with increasing concentrations of each of the three proteins. (b) The 250-bp *Ava*I-*Pst*I from the 5′ end of *eve* was ^{32}P labeled at the *Ava*I site on the coding strand and incubated with the same concentrations of each protein preparation. Lanes 2–4 correspond to increasing amounts of the wild-type *eve* protein; lanes 6–8 show increasing amounts of the *en* protein; and lanes 10–12 show increasing amounts of the ΔPvu mutant *eve* protein. Each titration point includes a fivefold increase in the concentration of protein. For example, lane 3 contains fivefold more *eve* protein than does lane 2, etc. Increasing concentrations of the wild-type *eve* protein resulted in three regions of protection within the 5′ region of *en* (a), which correspond to the k1, k2, and k3 sites. Note that concentrations of the *eve* protein that filled the k1 and k2 sites also protected the e4 and e5 sites within the 5′ end of *eve* (b). The e4 site actually contains two half-sites (called e4^W and e4^S), which were not uniformly filled by the *eve* protein. The lowest concentration of the *eve* protein that was assayed (lane 2) nearly filled the e4^S half-site, but failed to fill the e4^W half-site and instead resulted in the appearance of hypersensitive bands in this region. *en* protein gave full protection of the k1 and k2 sites within the 5′ end of *en* (a) but failed to bind detectably to e4 or e5 (compare lanes 8 of a and b). The binding of the *en* protein to k1, as judged by the appearance of hypersensitive bands, can be detected at the lowest protein concentration that was assayed (see lane 6 of a). A 25-fold-higher concentration of the protein failed to bind e4 or e5 (see lane 8 of b). The ΔPvu mutant *eve* protein failed to bind to either the 5′ *en* sites or 5′ *eve* sites. Lanes labeled C are controls done with 5 μg of extract prepared from IPTG-induced BL21(DE3) cells carrying the T7 expression vector without inserts. Lanes labeled − were incubated without protein, and lanes labeled GA are the deoxyguanosine and deoxyadenosine sequence of the labeled DNA strand.

Fig. 3, suggest that the *eve* homeo domain mediates binding to both classes of recognition sites and that the 3.77.17 protein might cause a disruption in the configuration of the protein so that binding to the class I sites is preferred over binding to the class II sites.

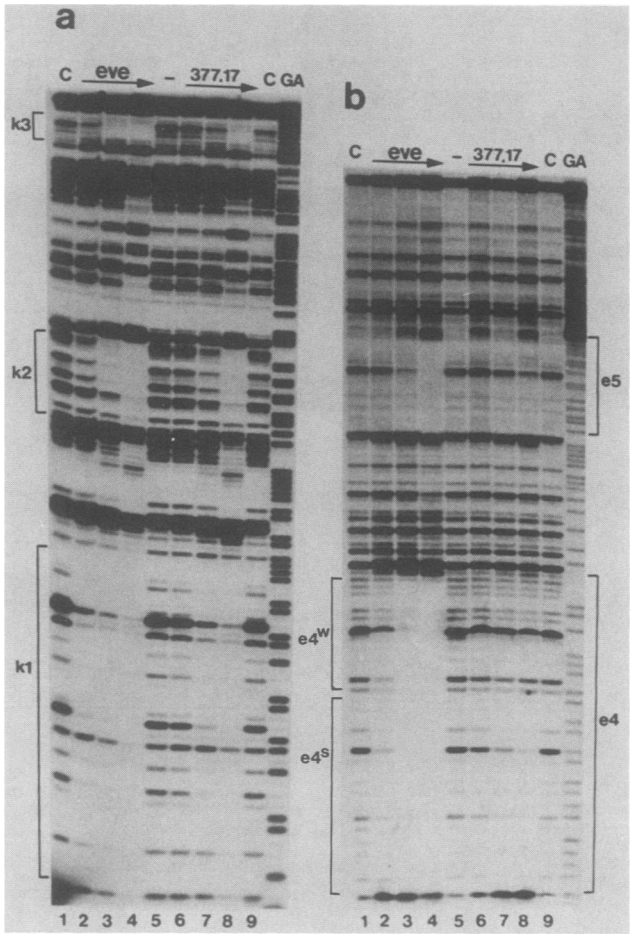

FIG. 4. Binding of the wild-type and 3.77.17 *eve* proteins. (a) The 5′ *en* *Cla*I-*Hin*fI fragment was labeled as described in the legend to Fig. 3. (b) The 5′ *eve* fragment was labeled at an *Hpa*II site located about 20 bp upstream from the e4 border. The samples in lanes 2 through 4 were incubated with increasing amounts of the wild-type *eve* protein extract, and lanes 6 through 8 were incubated with increasing amounts of the mutant 3.77.17 protein. As for Fig. 3, each titration point represents a fivefold increase in protein concentration. The binding of the 3.77.17 protein to the 5′ *en* k1, k2, and k3 sites was nearly indistinguishable from the binding of the wild-type *eve* protein. In contrast, the 3.77.17 protein showed a marked reduction in binding to the 5′ *eve* e4 and e5 sites compared with the wild-type protein. A concentration of the 3.77.17 protein that gave nearly full protection of the 5′ *en* k1 and k2 sites showed only weak binding to the 5′ *eve* e4 and e5 sites (see lanes 8). Note that the 3.77.17 protein was more impaired in its binding to the e4^W half-site than to the e4^S half-site.

Homeo box swap. Further evidence that the *eve* homeo domain mediates binding to both classes of recognition sequences stems from studies done with an *en* protein that contains the *eve* homeo domain and an *eve* protein that contains the *en* homeo domain. As discussed previously, the *eve* protein is unique in showing about equal preference for the class I and class II binding sites. Other homeo box proteins, such as *en*, display a much greater preference for the class I sites than for the class II sites (summarized in Table 1). For example, a concentration of the wild-type *en* protein that was sufficient to fully protect the 5′ *en* k2 site did not detectably bind to the 5′ *eve* e4 site (compare lane 8 of Fig. 3a with lane 8 of Fig. 3b). In order to determine whether the failure of the wild-type *en* protein to bind efficiently to

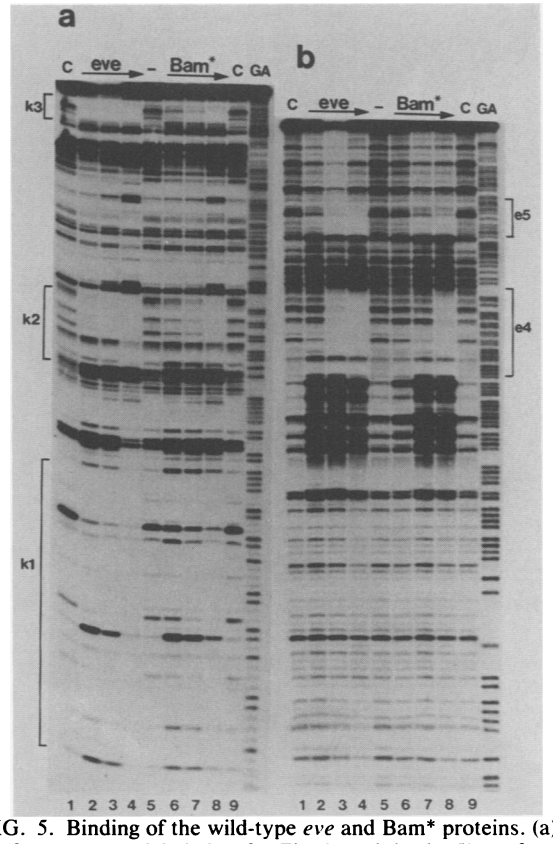

FIG. 5. Binding of the wild-type *eve* and Bam* proteins. (a) The 5′ *en* fragment was labeled as for Fig. 3, and (b) the 5′ *eve* fragment was labeled at the *Ava*I site on the noncoding strand. As for Fig. 3 and 4, each titration point represents a fivefold increase in protein concentration. Note that a concentration of the truncated Bam* protein that filled the 5′ *en* k1 and k2 sites (lane 8 of a) also filled the 5′ *eve* e4 site. Thus, the deletion of the *eve* carboxyl terminus does not significantly affect the preference of the protein for class I over class II recognition sequences. Compare lanes 3 and 8 in a and b.

the class II sites was a property of the *en* homeo domain, we examined the binding activity of a full-length *eve* protein that contained the *en* homeo domain. Conversely, in order to determine whether the efficient binding of the wild-type *eve* protein to both classes of recognition sequences was a property of the *eve* homeo domain, we examined an *en* protein that contains the *eve* homeo domain. The strategy used to prepare these recombinant *eve* and *en* proteins is outlined in Fig. 6. Mutagenic oligonucleotides were used to create *Xba*I and *Sma*I sites on either side of the *eve* and *en* homeo boxes (39). Note that the resulting *eve* and *en* coding sequences were not identical to the respective native proteins, in that several of the codons that reside within the newly created restriction sites were changed. However, despite these changes, the mutagenized *eve* and *en* proteins (called eve^{XS} and en^{XS}) showed normal DNA-binding activities. Increasing concentrations of the eve^{XS} and en^{XS} proteins filled the 5′ *en* k1, k2, and k3 sites with about the same kinetics as observed for the corresponding wild-type proteins (see lanes 1 through 8, Fig. 7a; compare with Fig. 3a). Moreover, concentrations of the eve^{XS} protein that protected the 5′ *en* k1 and k2 sites also filled the 5′ *eve* e4 site (compare lanes 3 and 4 of Fig. 7a with lanes 3 and 4 of Fig. 7b). The *Xba*I-*Sma*I homeo box fragments within the eve^{XS} and en^{XS} coding sequences were exchanged in order to prepare an *eve* protein with the *en* homeo domain (called

TABLE 1. Relative binding affinities of homeo box proteins[a]

Protein	Relative binding affinity	
	k1/e4^{S}	k1/e4^{W}
wt *eve*	1	3
wt *zen*	>25	>25
wt *en*	>25	>25
$eve^{3.77.17}$	2	10–20
eve-Bam*	1	3
eve-Acc*	1	3
eve^{XS}	1	3
en^{XS}	>25	>25
eve^{ENHB}	10	>10
en^{EVEHB}	10	>10

[a] The estimates are based on the amounts of each of the protein extracts required to give an equivalent level of protection within the 5′ *en* k1 site and each of the two 5′ *eve* e4 half-sites. The numbers represent the ratio of binding to k1 vs. that to e4 based on the footprint assays shown in Fig. 3, 4, 5, and 7. For example, the wild-type *eve* protein shows equal binding to k1 and e4^{S}, but shows a lower affinity for the e4^{W} half-site in that three times more of the extract is required to give a level of protection comparable to that observed for k1. When a protein fails to bind detectably to a site at the concentrations that were assayed, only a lower limit can be estimated. For example, even the highest concentrations of the wild-type *en* protein failed to protect either of the e4 half-sites, but a 25-fold-lower concentration of the protein provided detectable binding to k1. wt, Wild type.

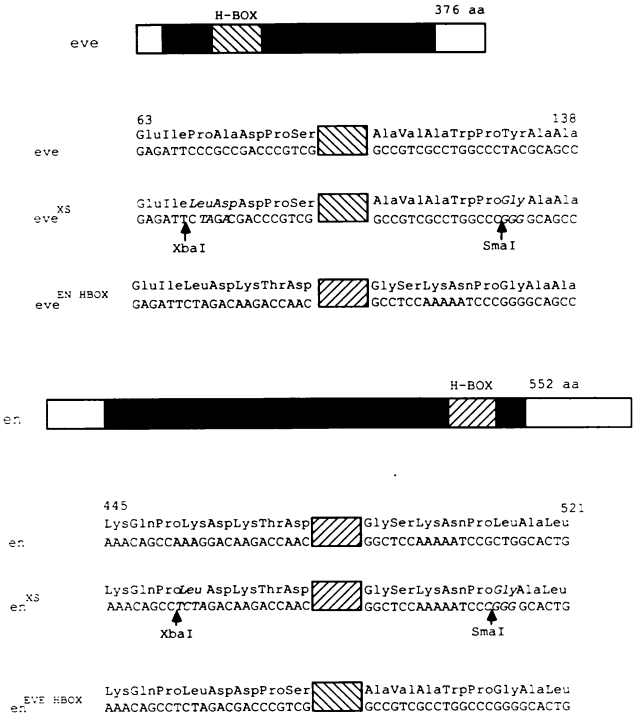

FIG. 6. Preparation of *eve* and *en* homeo box swap proteins. *Xba*I and *Sma*I sites were created on either side of the *eve* and *en* homeo boxes by oligonucleotide-directed mutagenesis. The nucleotide and corresponding amino acid (aa) substitutions resulting from the mutagenesis in the eve^{XS} and en^{XS} coding sequences are indicated by italics. The *Xba*I-*Sma*I fragments from the two modified coding sequences were exchanged to make the eve^{ENHB} and en^{EVEHB} recombinant proteins. As in Fig. 2a, the shaded rectangles indicate the coding regions, the hatched regions indicate the locations of the homeo box, and the open rectangles represent untranslated regions.

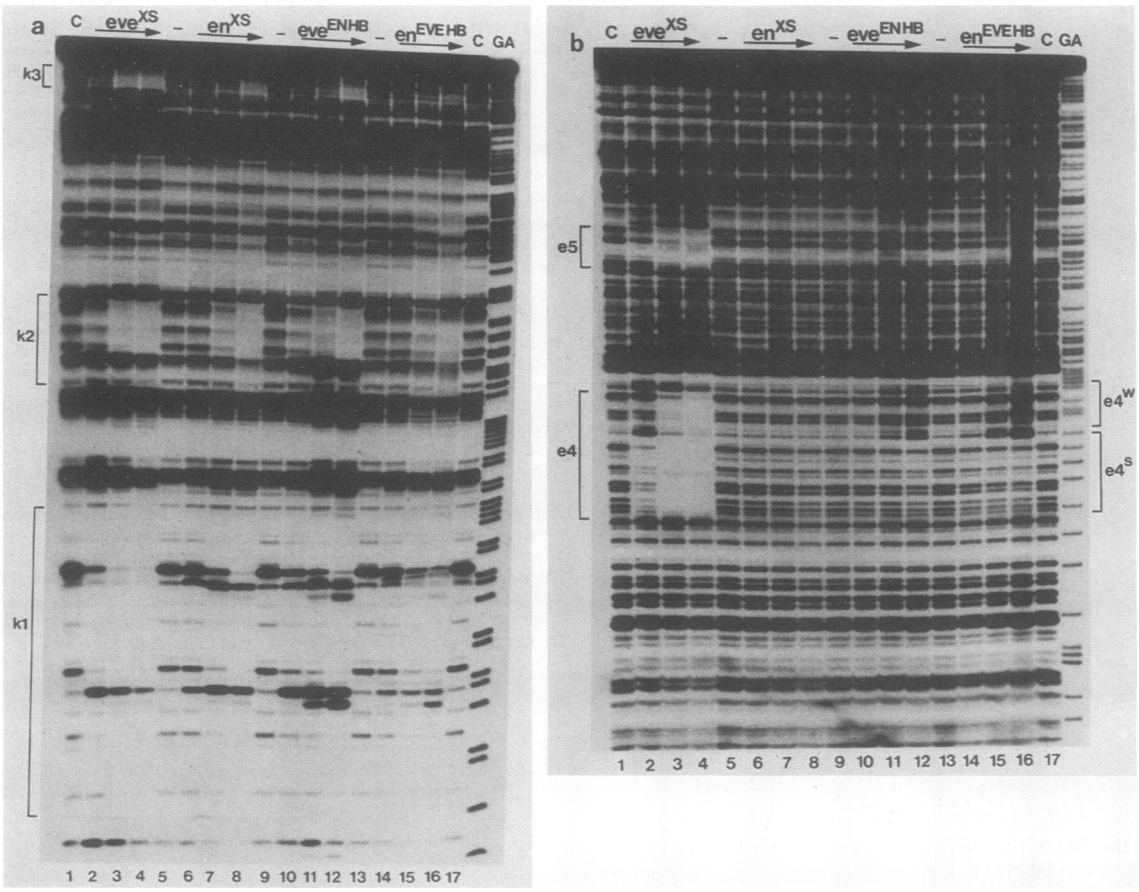

FIG. 7. Binding of the *eve* and *en* homeo box swap proteins. (a) The 5′ *en* fragment; (b) the 5′ *eve* fragment (labeled as in Fig. 3). In this experiment each titration point represents a threefold increase in protein concentration. Lanes 2–4, eve^{XS} protein; lanes 6–8, en^{XS} protein; lanes 10–12, *eve* protein with the *en* homeo box (eve^{XS} and en^{XS} proteins displayed the same binding properties as the corresponding wild-type *eve* and *en* proteins. The eve^{XS} protein showed about the same affinity for the 5′ *en* k1 site and the 5′ *eve* e4 site (see lanes 3 and 4 of a and b). The en^{XS} protein showed the same affinity to the 5′ *en* k1, k2, and k3 sites as did the wild-type *en* protein (compare lane 8 of a with lane 8 of Fig. 3a). Neither the wild-type *en* protein nor en^{XS} showed detectable binding to the 5′ *eve* sites (see lane 8 of b). Both homeo box swap proteins (eve^{ENHB} and en^{EVEHB}) showed DNA-binding properties that were intermediate between those of the wild-type *eve* and wild-type *en* proteins. The eve^{ENHB} protein (lanes 10–12) showed a strong preference for the 5′ *en* k1 site over the 5′ *eve* e4 site, but unlike the wild-type *en* protein it detectably bound to e4. Hypersensitive bands can be observed in the $e4^{W}$ half-site (see lanes 11 and 12 of b), indicating weak binding of the protein to $e4^{W}$. Moreover, the highest concentration of the protein showed nearly half-maximal protection of the $e4^{S}$ half-site (compare lanes 12 and 13 of b). The en^{EVEHB} protein (lanes 14–16) showed a binding behavior similar to that observed for the eve^{ENHB} protein.

eve^{ENHB}) and an *en* protein with the *eve* homeo domain (called en^{EVEHB}), as indicated in Fig. 6. Note that the recombinant proteins did not represent a precise exchange of *eve* and *en* homeo boxes, since the *Xba*I and *Sma*I sites occurred five codons on either side of each homeo box. The eve^{ENHB} and en^{EVEHB} proteins displayed binding activities that were quite distinct from those of the corresponding eve^{XS} and en^{XS} proteins. A concentration of the eve^{ENHB} protein that provided full protection of the 5′ *en* k2 site showed only weak binding to the 5′ *eve* e4 site (compare lane 12, Fig. 7a, with lane 12, Fig. 7b), whereas both the wild-type *eve* and eve^{XS} proteins bound to these sites with about equal affinity. Although greatly reduced, the eve^{ENHB} protein exhibited a low affinity for the e4 site. The weak half-site became progressively more hypersensitive with increasing amounts of the recombinant protein, similar to that observed with low concentrations of the wild-type *eve* protein. Furthermore, the highest concentration of the protein that was assayed provided about half-maximal protection of the strong half-site (compare lanes 12 and 13 of Fig. 7b). Thus, the eve^{ENHB} protein displayed a binding behavior that was intermediate between those of wild-type *eve* and *en* proteins. This suggests that the *en* homeo domain is responsible, at least in part, for the strong preference that the wild-type *en* protein displays for the class I over the class II binding sites. However, the weak affinity that the eve^{ENHB} protein displayed for the e4 site suggests that the new protein context of the *en* homeo domain increases its ability to mediate binding to the class II recognition sites (see Discussion).

In the reciprocal experiment, an *en* protein that contained the *eve* homeo domain (called en^{EVEHB}) was tested for DNA-binding specificity. The highest concentrations of the en^{EVEHB} protein that were assayed fully protected the 5′ *en* k1 and k2 sites (see lane 16 of Fig. 7a). With binding to the k1 and k2 sites used as internal standards, this concentration of the en^{EVEHB} protein was comparable to only the second titration point of the wild-type and eve^{XS} proteins that were used in this study (compare lanes 3 and 16 of Fig. 7a). This concentration of the wild-type or eve^{XS} protein was sufficient to completely fill the strong half-site within the 5′ *eve* e4 region and nearly filled the weak half-site. A comparison

of lanes 3 and 16 of Fig. 7 shows that the en^{EVEHB} protein gave significant protection within the e4 region, although its relative affinity for e4 was somewhat lower than that of the wild-type *eve* protein. The en^{EVEHB} protein gave nearly half-maximal protection of the strong e4 half-site and resulted in the appearance of several hypersensitive bands within the weak half-site (compare lanes 16 and 17, Fig. 7b). These results suggest that the *eve* homeo domain can mediate binding to both class I and class II recognition sequences, even in the context of the *en* protein.

DISCUSSION

A full-length *eve* protein prepared in bacteria binds with about equal affinity to two different classes of recognition sequences: A+T-rich class I sites that are found at the 5′ end of the *en* gene and G+C-rich class II sites that occur within the *eve* promoter. In vivo circuitry studies suggest that eve^{+} gene activity is required for normal patterns of *en* and *eve* expression (12, 24; Frasch and Levine, in preparation). The DNA-binding studies described here and in a previous report (15) are consistent with the possibility that one or both of these regulatory interactions occur at the level of transcription. The *eve* protein is unique among the homeo box proteins that have been examined in showing equal preference for the class I and class II binding sites. At least three of the other four proteins that have been examined display a strong preference for the class I sites over the class II sites (summarized in Table 1). Here we have demonstrated that the binding of the *eve* protein to the two classes of recognition classes can be uncoupled. Although the *eve* homeo domain appears to be responsible for both binding activities, the protein context of the homeo domain can influence its preference for the class I versus class II sites. The uncoupling of the two binding activities that was observed for the $eve^{3.77.17}$ mutant protein in vitro correlates well with the uncoupling of *eve* and *en* expression that was observed for 3.77.17 mutants in vivo and lends further evidence that *eve* regulates *eve* and *en* expression at the level of transcription.

Does the 3.77.17 protein disrupt a cooperative binding process? Tables 1 and 2 summarize the relative affinities of different *eve* and *en* mutant proteins for class I and class II binding sites. Increasing concentrations of each of the proteins examined in this study gave full protection of the 5′ *en* k1 and k2 sites, which served as internal standards for determining relative binding to class II sequences, such as the 5′ *eve* e4 site. e4 contains two half-sites that displayed slightly different affinities for the wild-type *eve* protein and were designated $e4^{S}$ (strong half-site) and $e4^{W}$ (weak half-site). The wild-type *eve* protein bound equally well to the k1 site and $e4^{S}$ half-site but displayed a threefold-lower affinity for $e4^{W}$ (Table 1). The 3.77.17 protein, which contained 79 extraneous amino acid residues in place of the normal carboxyl-terminal one-third of the native protein, showed a significant relative reduction in binding to the e4 site. There was at least a twofold relative reduction in binding to $e4^{S}$ and an even stronger reduction in binding to the $e4^{W}$ half-site (Table 2). One explanation for this reduced binding to the class II sequences is that the carboxyl-terminal one-third of the native *eve* protein specifies a second DNA-binding activity that is separate from the homeo domain. Perhaps the *eve* homeo domain mediates binding to class I sequences, whereas the carboxyl terminus mediates binding to class II sites. Several lines of evidence are inconsistent with this possibility. First, there is no precedent in either procaryotes or eucaryotes for a DNA-binding protein that contains more than one binding domain. Second, truncated *eve* proteins, such as Bam* and AccI* (Fig. 2, and Tables 1 and 2), displayed equally strong binding to both class I and class II recognition sequences. Third, a carboxyl-terminal peptide (ΔNdeI, Fig. 2) did not display any DNA-binding activity in our assays. And finally, a small in-frame deletion within the homeo domain (ΔPvuII, Fig. 2 and 3) completely abolished both binding activities.

TABLE 2. Normalized binding activities of mutant proteins[a]

Protein	Relative binding activity	
	$e4^{S}$	$e4^{W}$
$eve^{3.77.17}$	0.5	0.2
eve-Bam*	1	1
eve-Acc*	1	1
eve^{XS}	1	1
en^{XS}	—	—
eve^{ENHB}	0.1	>0.1
en^{EVEHB}	0.1	>0.1

[a] The activities of the different protein extracts were normalized based on binding to the 5′ *en* k1 site. The values correspond to the relative binding of normalized wild type (wt) *eve* extracts divided by that of each of the indicated mutant proteins to the e4 half-sites.

An alternative explanation for the binding behavior of the 3.77.17 protein is that the extraneous amino acid residues which replace the normal carboxyl terminus alter the overall configuration of the protein, impairing its normal binding activity. Perhaps the binding of the wild-type *eve* protein to $e4^{S}$ promotes binding to $e4^{W}$ through a cooperative process. This could involve protein-protein interactions or a distinct process by which the binding of the *eve* protein to $e4^{S}$ changes the topology of the $e4^{W}$ site so that it is more easily filled by the *eve* protein. There is a precedent for this latter type of mechanism: the binding of the *Drosophila* heat shock transcription factor causes a bend in the DNA and facilitates binding of the protein to an adjacent site (34). The 3.77.17 protein is far more impaired in its ability to bind to $e4^{W}$ than to $e4^{S}$ (Tables 1 and 2). The binding of the mutant protein to $e4^{S}$ might fail to "expose" the $e4^{W}$ site for efficient binding. Consistent with this interpretation is the finding that the *eve* homeo domain failed to promote binding of a recombinant *en* protein (en^{EVEHB}) to $e4^{W}$, even though this protein bound reasonably well to the $e4^{S}$ site (Tables 1 and 2). Since the truncated Bam* and AccI* *eve* proteins bound with normal affinity to both e4 half-sites (Tables 1 and 2), it would appear that the region of the wild-type *eve* protein that confers efficient binding to $e4^{W}$ resides within the amino-terminal half of the protein and might include the homeo domain.

Relative contributions of the homeo domain and protein context. The DNA-binding activities exhibited by the recombinant eve^{ENHB} and en^{EVEHB} proteins (Fig. 7) suggest that both the homeo domain and its protein context play a role in the efficient binding of the wild-type *eve* protein to the class II recognition sequences. The wild-type *en* protein shows a stronger preference for the class I binding sites than for the class II sites (Table 1), which suggests that the *en* homeo domain possesses little affinity for class II sequences. The *eve* and *en* homeo domains are quite divergent and share less than 50% amino acid identity (8, 24, 29), and this divergence might account for fundamental differences in their binding preferences. Even the putative recognition helices of the *eve* and *en* homeo domains are different, and although this corresponds to a highly conserved region of most homeo box genes (reviewed in reference 11), they are identical in only six of the nine amino acid residues that constitute this

region. However, the demonstration that the eve^{ENHB} protein bound the $e4^{S}$ site suggests that the context of the *en* homeo domain can influence its DNA-binding properties (Tables 1 and 2). The *eve* and *en* proteins show no significant homologies outside their respective homeo domains, and once the *en* homeo domain is placed in the context of an otherwise normal *eve* protein, it can mediate binding to $e4^{S}$. In contrast, the *en* homeo domain does not promote binding to e4 in the context of the *en* protein. However, the *en* homeo domain is not as efficient as the *eve* homeo domain in mediating binding to e4. The eve^{ENHB} protein failed to protect the $e4^{W}$ site and showed a 10-fold-lower relative binding to the $e4^{S}$ site than the wild-type *eve* protein. This suggests that sequence divergence between the *eve* and *en* homeo domains is at least partly responsible for the different binding preferences exhibited by the wild-type *eve* and *en* proteins.

The role of protein context in selecting binding preference is also evident from the binding behavior of the en^{EVEHB} recombinant protein. This protein showed a greater preference for the class II sequences than did the wild-type *en* protein (Table 1). This increase in relative binding to the $e4^{S}$ site indicates the important role that the *eve* homeo domain plays in mediating the binding of the wild-type *eve* protein to both class I and class II recognition sequences. However, the *eve* homeo domain is not sufficient to confer the same efficient binding to the class II sites as observed for the wild-type *eve* protein. In fact, the en^{EVEHB} protein exhibited about the same relative binding to the $e4^{S}$ site as observed for the eve^{ENHB} protein (Table 2). Both proteins displayed less efficient binding to the class II sites (particularly $e4^{W}$) than the wild-type *eve* protein in binding class II sites. These observations suggest that the *eve* homeo domain and nonhomeo regions of the wild-type *eve* protein play about equivalent roles in promoting efficient binding to class II recognition sequences. Thus, it would appear that the general failure of other homeo box proteins such as *en* and *zen* to bind class II sites (15) is a property of both their homeo domains and the flanking regions unique to these proteins. Binding to the class I sites, which is a feature common to at least six different homeo box proteins (4, 15; T. Hoey, unpublished results), is more directly dependent on the homeo domain and does not significantly involve protein context. Consistent with this conclusion is the finding that *en* fusion proteins which include little more than the *en* homeo domain bind to the 5′ *en* sites about as efficiently as does a full-length *en* protein (4; T. Hoey and C. Desplan, unpublished results).

***eve* autoregulation and the DNA-binding behavior of the 3.77.17 protein.** The DNA-binding activity of the 3.77.17 protein correlates quite closely with the uncoupling of the *eve* and *en* expression patterns that is observed in 3.77.17 mutants in vivo. eve^{+} gene activity is required for the establishment of both the *eve* and *en* expression patterns during early development (6, 12, 24; Frasch and Levine, in preparation). Null mutants cause partial pairwise fusions between adjacent *eve* stripes, and there is a failure to activate *en* expression in middle-body regions. Weaker *eve* mutations, such as 3.77.17, result in abnormally broad bands of *eve* expression during later stages of development and a nearly normal activation of *en* expression. The *en* pattern is slightly altered in that there are partial pairwise fusions between adjacent *en* stripes, which coincides with the altered *eve* pattern observed in these mutants (Frasch and Levine, in preparation). Despite this relatively minor disruption of the *en* pattern, the timing and levels of *en* expression are essentially normal. These observations suggest that the 3.77.17 mutation uncouples *eve* autoregulation and the regulation of *en*. Here we have shown that the 3.77.17 protein exhibits normal binding to the 5′ *en* sites, which correlates with the essentially normal expression of the *en* gene observed in these mutants in vivo. In contrast, the protein shows a substantially reduced binding to 5′ *eve* sites, which correlates with the abnormal *eve* pattern that is observed. These results lend further support to the proposal that the *eve* protein regulates *en* and *eve* expression in vivo at the level of transcription. Perhaps *eve* null proteins, such as that encoded by the temperature-sensitive allele eve^{ID19}, are unable to bind to both the 5′ *eve* and 5′ *en* sites, thereby accounting for the failure to correctly regulate either of these genes in such mutants. Consistent with this possibility is the finding that the temperature-sensitive allele maps within the *eve* homeo domain, very close to the putative recognition helix (J. Tugwood and M. Levine, unpublished results).

Future experiments will involve the use of P-element-mediated germ line transfer to express some of the mutant proteins described in this study in wild-type and mutant embryos. For example, to what extent can recombinant eve^{ENHB} or en^{EVEHB} proteins complement *eve* mutant embryos? Such studies might help resolve the role of the homeo domain and nonhomeo regions in executing the in vivo activities of *Drosophila* homeo box genes.

ACKNOWLEDGMENTS

We thank Claude Desplan and Pat O'Farrell for sharing unpublished results and for their encouragement.

This work was supported by a grant provided by the National Institutes of Health.

LITERATURE CITED

1. **Akam, M.** 1983. The location of *Ultrabithorax* transcripts in *Drosophila* tissue sections. EMBO J. **2**:2075–2084.
2. **Akam, M.** 1987. The molecular basis for metameric pattern in the *Drosphila* embryo. Development **101**:1–22.
3. **Carroll, S. B., and M. P. Scott.** 1986. Zygotically-active genes that affect the spatial expression of the *fushi tarazu* segmentation gene during early *Drosophila* embryogenesis. Cell **45**:113–126.
4. **Desplan, C., J. Theis, and P. H. O'Farrell.** 1985. The *Drosophila* developmental gene, *engrailed*, encodes a sequence-specific DNA binding activity. Nature (London) **318**:630–635.
5. **DiNardo, S., J. M. Kuner, J. Theis, and P. H. O'Farrell.** 1985. Development of embryonic pattern as revealed by accumulation of the nuclear *engrailed* protein. Cell **43**:59–69.
6. **DiNardo, S., and P. H. O'Farrell.** 1987. Establishment and refinement of segmental pattern in the *Drosophila* embryo: spatial control of *engrailed* expression by pair-rule genes. Genes Dev. **1**:1212–1225.
7. **Fjose, A., W. McGinnis, and W. J. Gehring.** 1985. Isolation of a homeo box-containing gene from the *engrailed* region of *Drosophila* and the spatial distribution of its transcripts. Nature (London) **313**:284–289.
8. **Frasch, M., T. Hoey, C. Rushlow, H. Doyle, and M. Levine.** 1987. Characterization and localization of the *even-skipped* protein of *Drosophila*. EMBO J. **6**:749–759.
9. **Frigerio, G., M. Burri, D. Bopp, S. Baumgartner, and M. Noll.** 1986. Structure of the segmentation gene *paired* and the *Drosophila* PRD gene set as part of a gene network. Cell **47**:735–746.
10. **Garcia-Bellido, A.** 1977. Homeotic and atavic mutations in insects. Am. Zool. **17**:613–629.
11. **Gehring, W. J.** 1987. Homeo boxes in the study of development. Science **236**:1245–1252.
12. **Harding, K., C. Rushlow, H. J. Doyle, T. Hoey, and M. Levine.** 1986. Cross-regulatory interactions among pair-rule genes in *Drosophila*. Science **233**:953–959.

13. **Harding, K., C. Wedeen, W. McGinnis, and M. Levine.** 1985. Spatially regulated expression of homeotic gene expression in *Drosophila*. Science **229**:1236–1242.
14. **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.
15. **Hoey, T., and M. Levine.** 1988. Divergent homeo box proteins recognize similar DNA sequences in *Drosophila*. Nature (London) **332**:858–861.
16. **Ingham, P. W., N. E. Baker, and A. Martinez-Arias.** 1988. Regulation of segment polarity genes in the *Drosophila* blastoderm by *fushi tarazu* and *even-skipped*. Nature (London) **331**: 73–75.
17. **Ingham, P. W., and A. Martinez-Arias.** 1986. The correct activation of antennapedia and bithorax complex genes requires the *fushi tarazu* gene. Nature (London) **324**:592–597.
18. **Kilcherr, F., S. Baumgartner, D. Bopp, and M. Noll.** 1986. Isolation of the *paired* gene of *Drosophila* and its spatial expression during early embryogenesis. Cell **45**:493–499.
19. **Kornberg, T., I. Siden, P. H. O'Farrell, and M. Simon.** 1985. The *engrailed* locus of *Drosophila*: *in situ* localization of transcripts reveals compartment-specific expression. Cell **42**:309–316.
20. **Laughon, A., and M. P. Scott.** 1984. Sequence of a *Drosophila* segmentation gene: protein structure homology with DNA-binding proteins. Nature (London) **310**:25–31.
21. **Lawrence, P. A., P. Johnston, P. Macdonald, and G. Struhl.** 1987. Borders of parasegments are delimited by the *fushi tarazu* and *even-skipped* genes. Nature (London) **328**:440–442.
22. **Levine, M., E. Hafen, R. Garber, and W. J. Gehring.** 1983. Spatial distribution of *Antennapedia* transcripts during *Drosophila* development. EMBO J. **2**:2037–2046.
23. **Levine, M., and K. Harding.** 1987. Spatial regulation of homeo box gene expression in *Drosophila* p. 116–142. *In* N. Maclean (ed.), Oxford survey on eukaryotic genes, vol. 5. Oxford University Press, Oxford.
24. **Macdonald, P. M., P. W. Ingham, and G. Struhl.** 1986. Isolation, structure, and expression of *even-skipped*: a second pair-rule of *Drosophila* containing a homeo box. Cell **47**:721–734.
25. **Maxam, A., and W. Gilbert.** 1980. Sequencing end-labeled DNA. Methods Enzymol. **65**:499–560.
26. **McGinnis, W., M. S. Levine, E. Hafen, A. Kuroiwa, and W. J. Gehring.** 1984. A conserved DNA sequence in homeotic genes of the *Drosophila* Antennapedia and Bithorax gene complexes. Nature (London) **308**:428–433.
27. **Nusslein-Volhard, C., H. Kluding, and G. Jurgens.** 1985. Genes affecting the segmental organization of the *Drosophila* embryo. Cold Spring Harbor Symp. Quant. Biol. **50**:145–154.
28. **Nusslein-Volhard, C., and E. Wieschaus.** 1980. Mutations affecting segment number and polarity in *Drosophila*. Nature (London) **287**:795–801.
29. **Poole, S. J., L. M. Kauvar, B. Drees, and T. Kornberg.** 1985. The *engrailed* locus of *Drosophila*: structural analysis of an embryonic transcript. Cell **40**:37–43.
30. **Rosenberg, A. H., B. N. Lade, D.-S. Chui, S.-W. Lin, J. J. Dunn, and F. W. Studier.** 1987. Vectors for selective expression of cloned DNAs by T7 RNA polymerase. Gene **56**:125–135.
31. **Rushlow, C., H. Doyle, T. Hoey, and M. Levine.** 1987. Molecular characterization of the *zerknullt* region of the Antennapedia gene complex in *Drosophila*. Genes Dev. **1**:1268–1279.
32. **Rushlow, C., K. Harding, and M. Levine.** 1987. Hierarchical interactions among pattern forming genes in *Drosophila*. Developmental toxicology mechanisms and risks, Banbury report 26. Cold Spring Harbor Laboratory, Cold Spring Harbor, N.Y.
33. **Scott, M. P., and S. B. Carroll.** 1987. The segmentation and homeotic gene network in early *Drosophila* development. Cell **51**:689–698.
34. **Shuey, D. J., and C. S. Parker.** 1986. Bending of promoter DNA on binding of heat shock transcription factor. Nature (London) **323**:459–461.
35. **Struhl, G., and R. A. H. White.** 1985. Regulation of the *Ultrabithorax* gene of *Drosophila* by other bithorax complex genes. Cell **43**:507–516.
36. **Studier, F. W., and B. A. Moffatt.** 1986. Use of bacteriophage T7 RNA polymerase to direct selective high-level expression of cloned genes. J. Mol. Biol. **189**:113–130.
37. **Wakimoto, B. T., F. R. Turner, and T. C. Kaufman.** 1984. Defects in embryogenesis in mutants associated with the *Antennapedia* gene complex of *Drosophila melanogaster*. Dev. Biol. **102**:147–172.
38. **Wedeen, C., K. Harding, and M. Levine.** 1986. Spatial regulation of antennapedia and bithorax gene expression by the *Polycomb* locus in *Drosophila*. Cell **44**:739–748.
39. **Zoller, M. J., and M. Smith.** 1984. Oligonucleotide-directed mutagenesis: a simple method using two oligonucleotide primers and a single-stranded DNA template. DNA **3**:479–488.