Identification and characterization of a gene activated by the deformed homeoprotein

James W. Mahaffey*, Denise F. Jones, Jo Ann Hickel and Charles M. Griswold
Department of Genetics, North Carolina State University, Raleigh NC 27695, USA

*Author for correspondence

SUMMARY

In Drosophila, the homeotic genes encode transcription factors which control segment identity during embryogenesis by specifying the appropriate set of ‘target’ genes necessary to produce the individual segmental characteristics. Though we know much about the homeotic genes and the proteins they encode, we know little of their targets. Here we identify and characterize one such target gene, a gene activated by the product of the homeotic gene Deformed. DNA binding assays and expression of reporter gene constructs indicate that activation of this gene requires a direct interaction between the Deformed protein and an upstream enhancer element at this target gene. However, although Deformed is required to activate this gene in cells of the maxillary segment, ectopically expressed Deformed does not activate the gene in other regions of the embryo. We conclude from this and other observations that additional factors may be required to activate the target gene, and, therefore, Deformed may participate in either a combinatorial or hierarchical activation signal in the maxillary cells. This newly identified gene encodes a novel protein of unknown function, though proteins with similar amino acid composition have been found. The pattern of transcript accumulation during embryogenesis indicates that this gene may be regulated by other homeoproteins in addition to Deformed.

Key words: Drosophila, homeotic genes, Deformed, target genes

INTRODUCTION

Study of the fruit fly Drosophila melanogaster has revealed a genetic hierarchy governing morphological development of the embryo (Nüsslein-Volhard and Wieschaus, 1980; reviewed by Akam, 1987; Ingham, 1988). Maternal gene products placed into the egg during oogenesis initiate the process by establishing the global embryonic axes and triggering expression of the zygotic segmentation genes. These segmentation genes subdivide the embryo into the familiar segmented insect body plan and, together with the earlier axial gene products, activate the appropriate homeotic gene(s) within each segment. The homeotic genes establish the appropriate segment identity both during embryonic and imaginal development (Lewis, 1978; Sanchez-Herrero et al., 1985; Mahaffey and Kaufman, 1988; Kaufman et al., 1990).

Homeotic genes encode determinants that act as switches selecting the developmental pathway a particular segment will follow (Lewis, 1978). Loss of a particular homeotic function causes segment identity defects in the embryo with the affected segment being transformed into that of a neighboring segment. Often, these segmental transformations are visible in the resulting mutant larval cuticle as duplications of certain pattern elements with a concomitant loss of others (Lewis, 1978; Wakimoto and Kaufman, 1981; Hayes et al., Sato et al., 1985). Likewise, ectopic expression of a homeotic gene is detrimental, altering the fate of segments that inappropriately accumulate the determinant (Lewis, 1978; Laughon et al., 1986; Frischer et al., 1986; Gibson and Gehring, 1988; Kuziora and McGinnis, 1988; Gonzalez-Reyes and Morata, 1990; Chadwick et al., 1990).

Proteins encoded by the homeotic genes (referred to as homeoproteins below) are transcription factors that control segment identity by selecting the appropriate battery of ‘target’ genes required to produce the segment-specific characteristics (Garcia-Bellido, 1977; Hayashi and Scott, 1990; Andrew and Scott, 1992). Homeoproteins contain a highly conserved, 60 amino acid peptide domain known as the homeodomain (McGinnis et al., 1984a,b; Laughon and Scott, 1984), a peptide motif conferring sequence-specific DNA binding ability to the homeoproteins (Laughon and Scott, 1984; Desplan et al., 1985; Beachy et al., 1988; Müller et al., 1988; Ekker et al., 1991). Since the initial discovery of the homeodomain, many proteins have been identified which contain the homeodomain sequence; several are known transcription factors (Strum et al., 1988; Clerc et al., 1988). Further, it has been shown through transfection experiments in tissue culture cells and by in vitro transcription assays that homeodomain-containing proteins have the potential to regulate target genes, supporting the role of homeoproteins as transcription factors (Vincent et
al., 1990; Krasnow et al., 1989; Winslow et al., 1989; John-

Though much is known about the homeotic genes and the
proteins they encode, little is known of the genes that
they regulate because the identification of such target genes
has proved difficult (Morata and Struhl, 1990; Andrew and
Scott, 1992). Unlike individual homeotic genes, target
genes may have roles throughout the organism, so that the
expression pattern or phenotype of mutants may not imme-
diately suggest that a gene is a target of a particular home-
opic protein. However, understanding how homeoproteins
function and determining the roles of the next genes in the
developmental hierarchy are compelling reasons to search
for target genes, and recently, several searches have been
fruitful (Bienz and Tremml, 1988; Reuter et al., 1990;
Gould et al., 1990; Wagner-Bernholz et al., 1991; Graba et

We began a search to identify genes that are regulated
by homeoproteins during embryonic development. We have
concentrated on identifying genes that are regulated by
those homeoproteins encoded by the Antennapedia-Com-
plex, specifically, the genes that specify head segment iden-
tity. Here, we report the identification and characterization
of one such gene, a gene regulated by Deformed, the
homeotic gene required in the mandibular and maxillary
head segments of the Drosophila embryo (Merrill et al.,
1987; Regulski et al., 1987; Jack et al., 1988; Mahaffey et
al., 1989). Below we present evidence that a direct inter-
action with the Deformed homeoprotein is required to ac-
tivate this target gene in cells of the maxillary segment. The
gene is expressed in other regions of the embryo and is
likely regulated by other homeoproteins in these areas. This
newly identified gene encodes a previously unknown pro-
tein.

MATERIALS AND METHODS

Drosophila stocks and culturing

Drosophila cultures were maintained at 25°C on standard corn-
meal-molasses-agar medium. The enhancer trap stocks are
described in Bier et al. (1989) and the transformation stocks are
described in Robertson et al. (1988). All other genetic symbols
are described in Lindsley and Zimm (1992). (Note, using both
molecular and cytological methods, we examined all of the stocks
available from the Bloomington, IN Drosophila Stock Center that
were reported to carry deficiencies of the 42B region. We found
that none actually deleted this region. The listings with the stock
center have been changed.)

Generation of enhancer trap lines

Enhancer trap insertion lines were established using the fly stocks
and procedures of Bier et al. (1989). Once lines were established,
flies (about 15, mixture of males and females) were placed in a
collection apparatus similar to that described by Wieschaus and
Nüsslein-Volhard (1986) using 15×100 mm polystyrene tubes.
Embryos were collected on grape juice agar plates. The embryos
were transferred to wire mesh staining baskets similar to those
described in Bier et al. (1989). Embryos were washed, fixed and stained in this staining basket
following a modification of the staining procedure of Mahaffey and Kaufman (1987); the main modification was that, during the
fixing stage embryos were not shaken but gently agitated by lift-
ing the container in and out of the fixing solutions. For all treat-
ments including antibody staining reactions, 3 ml of each solu-
tion were placed into the Petri dish. The staining basket was blot-
ted on Kimwipes between solution changes to remove excess
liquids. A monoclonal anti-β-galactosidase (β-gal) antibody
(Promega) was used to detect enhancer trap expression. Embryos
were exposed to the antibody (1:1000 dilution) for 1.5 hours,
washed for 30 minutes and incubated with an HRP-labeled goat
anti-mouse secondary antibody (Cappel). The embryos were again
washed for 30 minutes then stained using DAB. The stained
embryos were photographed using a Zeiss Axioplan microscope.

Antibody staining procedures

Detection of β-gal in transgenic embryos followed the immuno-
ological procedures described in Mahaffey and Kaufman (1987).
Both monoclonal (Promega) and polyclonal (Cappel) anti-β-gal
antisera were used. The Deformed antiserum was described in
Mahaffey et al. (1989). Expression of hsDfd essentially followed
the procedure of Kuziora and McGinnis (1988). Embryos were
collected for 3 hours after which they were heat shocked (37°C)
for 1 hour then allowed to recover for 0-8 hours prior to staining
to detect Deformed and β-gal.

Cloning of DNA at the enhancer trap insertion

Isolation of DNA flanking the enhancer trap element followed pro-
cedures recommended by Bier et al. (1989). Genomic DNA was
extracted from adult flies (Ashburner, 1989), digested with EcoRI
and ligated under dilute conditions to favor circle formation. The
DNA was transformed into SURE cells (Stratagene) which were
selected for ampicillin resistant colonies. Plasmid DNA was
extracted from resistant colonies (Sambrook et al., 1989) and used
for in situ hybridization to polytene chromosomes (Ashburner,
1989) to verify that the cloned DNA was from the same position
as the enhancer trap element.

Genomic and cDNA cloning

Genomic DNA encompassing the enhancer trap tagged region was
obtained from lambda phage libraries (Maniatis et al., 1978;
Promega). The cDNA clone was isolated from a 3-12 hour embry-
onic cDNA library (Poole et al., 1985). Plaques were lifted onto
nylon membranes (MSI) and autoclaved for 10 minutes (gravity
cycle) to lyse the phage and immobilize the DNA. Phage DNA
on the membranes was hybridized with the plasmid-rescued DNA
which was labeled using the MultiTm DNA Labeling System
(Amersham) and [α-32 P]dCTP (NEN Research Products). Hybridizing plaques were isolated and subclones were generated
using standard recombinant DNA techniques (Sambrook et al.,
1989). DNA sequences were determined using the Sequenase
system (USB).

Deformed homeoprotein DNA binding assays

The Deformed expression plasmid, pARDfd, was a gift from W.
McGinnis and is described in Jack et al. (1988). Induction of
Deformed protein expression was performed as described in
Studier and Moffatt (1986). Isolation of protein, DNA-protein
binding and immunoprecipitation of DNA-protein complexes were
performed as in Regulski et al. (1991) using anti-Deformed anti-
bodies described in Mahaffey et al. (1989). The subclones of the
I.28 region were digested with appropriate restriction enzymes to
produce 5′ overhangs, which were then filled in with [α-32 P]dCTP
(Sambrook et al., 1989). Binding reactions were carried out in 50
µl of binding buffer containing about 1 fmole of 32P-labeled DNA,
0.25 µg Deformed-containing protein extract and 100 µg/ml
sheared salmon sperm DNA. Reactions were incubated on ice for
30 minutes. Following the incubation, 15 µl of prepared Staphy-
llococcus aureus cells were added and incubated on ice for an addi-
205

Targets of the Deformed homeotic gene

Cells were then spun down (1,000 g for 30 seconds) and the supernatant discarded. The DNA-protein complexes were washed twice (2 minutes each) in 200 µl binding buffer containing 10 µg/ml salmon sperm DNA and resuspended in 200 µl TE (10 mM Tris, pH 8.0, 1 mM EDTA). This was extracted twice with a 1:1 volume of phenol/chloroform followed by a chloroform extraction. The DNA was ethanol precipitated in the presence of 20 µg of tRNA, and rehydrated in 5 µl TE. Precipitated DNA was analyzed on a 1.5% agarose gel. The plasmid pBST 1.4, containing the Deformed auto-activation region, a gift of W. McGinnis, was used as a positive control. To demonstrate that Deformed was required for the immunoprecipitation, an extract was prepared from bacteria lacking the pARDfd plasmid.

Construction of the enhancer tester reporter gene and Drosophila transformation

The plasmid pHZ-white, a gift from J. Gindhart and T. Kaufman, is a mini-white version of pHZ50 (Hiromi and Gehring, 1987) and was used to test for enhancer activity. First, a Xhol linker was cloned into the BamHI site of the shuttle vector pHSS7 to obtain a vector suitable for cloning SalI fragments. This shuttle vector allows fragments to be removed with NotI ends permitting cloning of the fragment into the NotI site of pHZ-white. The 4.0 kb SalI fragment was cloned into pHZ-white so that the genomic fragment inserted in the positive orientation with respect to the enhancer tester lacZ gene. Plasmids were extracted and purified by CsCl-ethidium bromide gradient centrifugation (Sambrook et al., 1989). The DNA was dialyzed against four changes of TE for eight hours, ethanol precipitated, and resuspended in 200 µl TE. To remove any residual CsCl, DNA was purified using an Elutip column (Schleicher and Schuell) and rehydrated in injection buffer (Spradling and Rubin, 1982; Ashburner, 1989) at a concentration of 400 ng/µl. Drosophila transformation followed the procedure of Robertson et al. (1988).

In situ localization of 1.28 transcripts

For the initial identification of the tagged gene, the genomic DNA fragment was labeled with digoxigenin (Boehringer Mannheim) using PCR. In situ hybridization to whole embryos followed the procedure of Tautz and Pfeifle (1989). Subsequent in situ analysis used ribonucleotide probes generated using an RNA Transcription Kit (Stratagene) and DIG-11-UTP (Boehringer Mannheim). Hybridization was carried out using modifications to the Tautz and Pfeifle (1989) method. Both horseradish peroxidase and alkaline phosphatase were used to detect hybridization.

RACE:PCR

PCR protocols essentially followed those of Frohman (1990) except that the first five amplifications were performed in the presence of only the dT-containing primer and at 32°C to facilitate annealing of the dT region to the poly(A) tract. The remaining primers were then added and cycling proceeded as described. All 1.28-specific primers were chosen from exon two to prevent product formation from any possible contaminating genomic sequence. PCR products were subcloned into pBluescript II KS + for sequencing.

RESULTS

β-gal expression in the 1.28 enhancer trap line suggests the tagged gene is a target of Deformed

We used a two step procedure beginning by identifying enhancer trap containing lines expressing the reporter gene in the embryonic head segments (O’Kane and Gehring, 1987; Bier et al., 1989; Wilson et al., 1989). The enhancer trap technique identifies genes by their pattern of expression. When the enhancer trap transposable element inserts within the vicinity of a gene, the reporter gene of the enhancer trap element (lacZ) responds to the transcriptional signals at the site of insertion, and therefore, β-gal expression mimics the pattern of expression of the endogenous gene. Since it is the enhancer that is tagged, we next looked for an effect on β-gal expression when the embryos containing the enhancer trap element were homozygous for
homeotic mutations. In one of our lines, line 1.28, the pattern of β-gal accumulation closely resembled that of the homeoprotein Deformed in the epidermis of the maxillary lobe. (Note, throughout this text we refer to the gene tagged by the enhancer trap line number 1.28. An appropriate name, reflecting any observed phenotype will be given to the gene when the genetic analysis is complete.) The embryo shown in Fig. 1 was stained to detect Deformed (Fig. 1A) and β-gal (Fig. 1B) using rhodamine- and fluorescein-labeled secondary antibodies, respectively. At this stage, Deformed accumulates within the posterior-lateral epidermal cells of the maxillary lobe, a few cells in the posterior mandibular lobe, within cells along the anterior edge of the dorsal ridge and within parasegment 1 of the ventral nerve cord (Jack et al., 1988; Mahaffey et al., 1989); in this figure we focus on the maxillary cells. Fig. 1C is a double exposure showing a one-to-one correspondence between β-gal and Deformed distribution within the epidermal cells of the maxillary lobe. This pattern of expression indicates that the gene tagged in the 1.28 line might be a target, either directly or indirectly, of the Deformed protein. We did not detect β-gal in other areas where Deformed was expressed; however, β-gal was detected in regions of the embryo not under Deformed control (proctodeum and hindgut, anterior and posterior midguts and in a small patch of cells in the posterior compartments of the first thoracic and eighth abdominal segments, see below).

**Deformed is required but not sufficient for 1.28 expression**

Next, we examined whether Deformed was required for β-gal expression in the maxillary cells. Since the 1.28 line is homozygous viable, we could generate a fly line homozygous for the enhancer trap insertion while heterozygous for Deformed<sup>w21</sup>, a loss-of-function mutation at Deformed (Merrill et al., 1987). One quarter of the embryos collected from this line were homozygous for the Deformed<sup>w21</sup> mutation while all embryos were homozygous for the 1.28 insertion. As shown in Fig. 2A, β-gal reporter expression was normal in embryos heterozygous for the Deformed<sup>w21</sup> mutation, while only background levels of β-gal were detected in the maxillary lobes of embryos homozygous for the Deformed<sup>w21</sup> mutation (Fig. 2B). Residual β-gal in the maxillary lobe did not resemble that seen in the original 1.28 line nor did it resemble the pattern of Deformed protein accumulation. β-gal distribution was not altered in any other regions of the embryo (Fig. 2C). In a related series of experiments, we examined β-gal distribution in embryos ectopically expressing Deformed in all tissues. These embryos were heterozygous for hsDfd (Kuziora and McGinnis, 1988) and for the 1.28 enhancer trap insertion. After administering the appropriate heat shock treatment required to activate the Deformed genes, and allowing for various recovery times, embryos were collected, fixed and stained to detect β-gal. We observed no effect on the β-gal distribution after ectopic Deformed expression (data not shown).

**Molecular cloning of the 1.28 gene**

The enhancer trap element allows rapid cloning of genomic DNA immediately flanking the insertion point (Bier et al., 1989). From 1.28 genomic DNA, a plasmid was recovered that included about 1.2 kb of DNA adjacent to the enhancer trap element. We used the rescued genomic DNA fragment as a probe to obtain lambda clones from *Drosophila* genomic libraries. In situ hybridization to polytene chromosomes using the rescued fragment and genomic clones as probes verified that all originated from the 42B region of chromosome 2 (Fig. 3). An abbreviated restriction map of the 1.28 genomic region is shown in Fig. 4.
TheTargets of the Deformed homeotic gene RNA of about 2.1 kb on a northern blot of 0-15 hour embryonic poly(A) + RNA (data not shown). In situ hybridization to whole embryos (Tautz and Pfeifle, 1989) demonstrated that the RNA detected by this fragment accumulated in a pattern identical to the β-gal reporter of the enhancer trap line, indicating that at least a portion of the tagged gene was located within this genomic interval (see below for in situ RNA localizations). This 6.5 kb fragment was also used as a probe to obtain cDNAs from a 3-12 hour cDNA library (Poole et al., 1985). A single cDNA was obtained from about 10^6 recombinants, and, although this cDNA is incomplete at both the 3′ and 5′ ends, it contains portions of two exons which are separated by an approx. 24 kb intron in the genomic DNA. Genomic DNA sequencing verified the splice junctions; both donor and acceptor splice consensus sequences were identified at the junctions of the intron. Genomic sequencing was also used to identify a potential transcriptional start site and poly(A) addition signal, and RACE:PCR (Frohman, 1990) was used to help identify the 5′ extent of the transcript and open reading frame. A potential start of transcription was found within an approx. 300 bp EcoRI fragment (Fig. 4). This sequence contains a match to the Drosophila start of transcription consensus as defined in Hultmark et al. (1986; Fig. 5) as well as other common eukaryotic transcription signals associated with promoters, namely TATA (−32) and GC (−59) boxes. A transcript beginning at this start site, spliced as observed in the cDNA, and ending at the poly(A) addition signal, would produce a transcript of 1999 nt. Polyadenylation of this transcript would bring it within the size range of the 2100 nt transcript observed in our northern analysis. Because of the unusual position of the termination codon (the first full codon in the second exon, see below) we verified the mRNA structure by sequencing additional cDNAs which were generated using PCR of reverse transcribed, oligo d(T)-primed embryonic mRNA. The PCR primers flanked the start of translation at the 5′ end and the splice junction at the 3′ end. The sequence of these cDNAs verified the coding region and the structure of the spliced mRNA.

**Sequence of transcript from the tagged gene and its potential protein**

The longest ATG-initiated open reading frame within the 1.28 transcribed region is 654 bp (Fig. 5). This open reading frame begins with a close match to the Drosophila translation start consensus sequence (Cavener, 1987) and ends with TAA. Interestingly, the termination codon is the first full codon of exon 2 (see splice site location in Fig. 5). Other ATG-initiated open reading frames within the 5′ portion of the transcript (12) are not within contexts which match the Drosophila start of translation consensus, and the longest open reading frame is 120 nt. With an open reading frame of 654 nt, the mRNA would have a 5′ untranslated region of 765 nt and a 3′ untranslated region of 580 nt. Translation of the open reading frame would produce a protein of about 22×10^3 Mr. The protein would be very basic (29 basic and 12 acidic residues) and have a high content of several amino acids, Ala (11.1%), Ser (11.5%), Thr

Fig. 3. Localization of the 1.28 gene on Drosophila polytene chromosomes. A shows the right arm of the second chromosome after orcein staining. B highlights an in situ hybridization using one of the 1.28 lambda clones as a probe.

RNA of about 2.1 kb on a northern blot of 0-15 hour embryonic poly(A)^+ RNA (data not shown). In situ hybridization to whole embryos (Tautz and Pfeifle, 1989) demonstrated that the RNA detected by this fragment accumulated in a pattern identical to the β-gal reporter of the enhancer trap line, indicating that at least a portion of the tagged gene was located within this genomic interval (see below for in situ RNA localizations). This 6.5 kb fragment was also used as a probe to obtain cDNAs from a 3-12 hour cDNA library (Poole et al., 1985). A single cDNA was obtained from about 10^6 recombinants, and, although this cDNA is incomplete at both the 3′ and 5′ ends, it contains portions of two exons which are separated by an approx. 24 kb intron in the genomic DNA. Genomic DNA sequencing verified the splice junctions; both donor and acceptor splice consensus sequences were identified at the junctions of the intron. Genomic sequencing was also used to identify a potential transcriptional start site and poly(A) addition signal, and RACE:PCR (Frohman, 1990) was used to help identify the 5′ extent of the transcript and open reading frame. A potential start of transcription was found within an approx. 300 bp EcoRI fragment (Fig. 4). This sequence contains a match to the Drosophila start of transcription consensus as defined in Hultmark et al. (1986; Fig. 5) as well as other common eukaryotic transcription signals associated with promoters, namely TATA (−32) and GC (−59) boxes. A transcript beginning at this start site, spliced as observed in the cDNA, and ending at the poly(A) addition signal, would produce a transcript of 1999 nt. Polyadenylation of this transcript would bring it within the size range of the 2100 nt transcript observed in our northern analysis. Because of the unusual position of the termination codon (the first full codon in the second exon, see below) we verified the mRNA structure by sequencing additional cDNAs which were generated using PCR of reverse transcribed, oligo d(T)-primed embryonic mRNA. The PCR primers flanked the start of translation at the 5′ end and the splice junction at the 3′ end. The sequence of these cDNAs verified the coding region and the structure of the spliced mRNA.

**Sequence of transcript from the tagged gene and its potential protein**

The longest ATG-initiated open reading frame within the 1.28 transcribed region is 654 bp (Fig. 5). This open reading frame begins with a close match to the Drosophila translation start consensus sequence (Cavener, 1987) and ends with TAA. Interestingly, the termination codon is the first full codon of exon 2 (see splice site location in Fig. 5). Other ATG-initiated open reading frames within the 5′ portion of the transcript (12) are not within contexts which match the Drosophila start of translation consensus, and the longest open reading frame is 120 nt. With an open reading frame of 654 nt, the mRNA would have a 5′ untranslated region of 765 nt and a 3′ untranslated region of 580 nt. Translation of the open reading frame would produce a protein of about 22×10^3 Mr. The protein would be very basic (29 basic and 12 acidic residues) and have a high content of several amino acids, Ala (11.1%), Ser (11.5%), Thr

Fig. 4. Genomic organization of the 1.28 region. Only selected restriction sites are shown. The start of transcription is 0 and size demarcations are shown every 5 kb. The position, orientation and features of the enhancer trap element are shown (P-lacW); the arrow above the element indicates the direction of transcription of the enhancer trap reporter lacZ. A line is drawn under the 6.5 kb XhoI-EcoRI fragment initially used as a probe to identify the transcribed region. The two exons (filled blocks) are separated by an approx. 24 kb intron; the intron has been truncated in the figure. Restriction sites are; H, HindIII; L, SalI; R, EcoRI; X, XhoI. The SalI site extending from the dotted line is a polymorphic site.
(13.4%), Pro (10.1%) and Gly (9.2%), some of which are not evenly distributed throughout the protein. For example, a proline-rich domain is located between amino acid positions 73-128 and glycine-rich domains are found in several locations, notably between positions 132-149. We found no significant homology to this translation sequence in searches of available databases using FASTA and BLAST programs; however, similar amino acid compositions are found in other proteins (see below).

1.28 expression in embryos

We used in situ hybridization to whole embryos to monitor 1.28 transcript distribution during embryogenesis. In all cases, transcript accumulation agreed with \( \beta \)-gal distribution, except \( \beta \)-gal expression was more stable, persisting throughout later developmental stages. In the maxillary lobe, 1.28 transcript is first detected during germ band contraction, mid to late stage 12 (Fig. 6D; embryo staging follows that of Campos-Ortega and Hartenstein, 1985). This is the same stage that we first detected \( \beta \)-gal in the maxillary lobes of embryos carrying the enhancer trap. As with \( \beta \)-gal, transcript accumulation is restricted to the posterior-lateral epidermis of the lobe. That Deformed is required for the maxillary expression of the 1.28 gene is shown in Fig. 6. Fig. 6A shows 1.28 transcript distribution in the maxillary segment of a wild type embryo that has just completed germ band contraction. Fig. 6B shows transcript localization in a similarly staged embryo obtained from Deformed \( w^{21} /+ \) parents; about one quarter of the embryos obtained from these parents lacked maxillary staining while
all other aspects of the hybridization signal were normal. We conclude that, as with β-gal expression from the enhancer trap, Deformed is required for expression of the 1.28 gene.

1.28 transcripts accumulate in other areas of the embryo, again agreeing with the pattern of β-gal expression. The earliest expression of 1.28, at the end of germ band contraction (stage 11), is observed in cells of the proctodeum at the posterior tip of the embryo, as well as within a cluster of cells in the posterior-lateral portion of the eighth abdominal segment (Fig. 6C,D). Transcripts also accumulate in the hindgut and in the developing anterior and posterior midguts (Fig. 6E). After germband contraction, we observed transcripts accumulating within a few cells along the posterior-mid-lateral edge of the first thoracic segment (Fig. 6F). Later in development, this group of cells forms a rounded structure slightly below the surface of this segment (Fig. 6G).

Though cytoplasmic staining was evident at low magnifications, at higher magnifications the in situ hybridization signal appeared as two distinct small dots within each nucleus (Fig. 6F), especially when horseradish peroxidase.

**Fig. 6.** Whole embryo in situ localization of the 1.28 transcript. (A) Transcript localization in the maxillary lobe (Mx) of an embryo that has completed germ band contraction. (B) A similar embryo from a cross of Deformed**w21**/+ parents. Note the lack of maxillary expression in the homozygous Deformed**w21** embryo. (C) Dorsal view of the posterior localization of 1.28 transcripts in an embryo just beginning germ band contraction. pr, proctodeum; A8, eighth abdominal segment. (D) First detection of the 1.28 transcript in the maxillary lobe of a mid-stage-12 embryo. The expression of 1.28 follows the changing aspects of Deformed expression at this stage. (E) Expression of 1.28 in the anterior (amg) and posterior (pmg) midguts and hindgut (hg) of a mid-stage-12 embryo. (F, G) 1.28 expression in posterior-lateral cells of the first thoracic segment (T1) within the region where the anterior spiracle will form (Mx, maxillary segment). F shows expression in an embryo beginning head involution while the embryo in G has almost completed this process. Note the recessed appearance of the signal in G.
histochemical staining was used. Shermoen and O’Farrell (1991) attributed these nuclear dots to the detection of nascent transcripts, detection of which may not be surprising for a gene containing an approx. 24 kb intron.

The Deformed protein binds to DNA upstream of the tagged gene

The pattern and timing of 1.28 expression along with the fact that Deformed is required to activate the 1.28 gene indicate that this gene is regulated by Deformed. If this regulation is direct, then an enhancer capable of specifying regulation by Deformed should be associated with the gene. This enhancer should have certain properties, among them the ability to be bound by the Deformed protein and the ability to direct maxillary expression of a reporter gene.

McGinnis and coworkers developed an antibody precipitation technique that can identify DNA fragments capable of being bound by the Deformed protein (Regulski et al., 1991). We used this same procedure to identify Deformed protein binding sites near the 1.28 gene. We tested DNA fragments throughout much of the 1.28 region (including the 1.28 cDNA) and, as a positive control, we used the Deformed auto-activation region (Bergson and McGinnis, 1990; Regulski et al., 1991). A 4.0 kb SalI DNA fragment, located at position −3500 to +500 with respect to the 1.28 gene, could be immunoprecipitated in this binding assay (Fig. 7). Immunoprecipitation of this fragment was dependent upon the presence of Deformed protein in the assay, as a bacterial extract lacking Deformed failed to precipitate the fragment. The 4.0 kb fragment contains several sequences closely matching the Deformed recognition sequences described by Regulski et al. (1991) as well as sequences matching the binding consensuses for several other transcription factors that specify early pattern formation (data not shown). The specificity of this procedure was demonstrated by the fact that the 300 bp EcoRI fragment, which includes the start of transcription, did not precipitate in this assay (data not shown). This fragment contains several closely spaced ATTA sequences (the core of the homeodomain binding sequence) 5’ to the start of transcription. It may be significant that only one of these ATTA sequences is in a context similar to the Deformed binding consensus.

The DNA fragment containing Deformed binding sites directs maxillary expression of enhancer tester constructs

That Deformed binding sites are present in the upstream DNA fragment indicates that a Deformed-responsive enhancer may be located there. To establish whether this fragment actually contains such an enhancer, we subcloned the 4.0 kb SalI fragment into the P element vector pHZ-white (the construct is referred to as HZ-4.0 below). This vector uses the E. coli lacZ gene as a reporter to test a chosen DNA fragment for enhancer activity. The vector is a derivative of the enhancer tester vector pHZ50 of Hiromi and Gehring (1987) except pHZ-white contains white as the selectable marker instead of the original ry. After establishing transformed stocks containing the HZ-4.0 construct, 0-22 hour embryos were collected and stained from each

Fig. 7. Deformed-DNA binding immunoprecipitation assay. A map of the upstream region of the 1.28 gene is at the top of the figure showing the location of the Deformed binding fragment. The start of transcription is at position no. 1. Restriction sites are as in Fig. 4. The dark bar indicates the extent of the SalI 4.0 kb fragment tested in the DNA binding assay (and subsequently in the enhancer tester experiments shown in Fig. 8). The 3’ SalI site of this 4.0 kb fragment (L) is an artificial site from the end of the genomic lambda clone. The position of the P element insertion is shown. The gel on the left shows the end-labeled DNA fragments used in the assay (1/10 the of the amount assayed), and includes the SalI 4.0 plasmid (digested with SalI) and pBST 1.4 (digested with HindIII) which contains the Deformed auto-activation region and was used as a control. The center gel shows the immunoprecipitated fragments. The panel on the right is a control using the same SalI 4.0 reaction but comparing the immunoprecipitation capability of an extract from bacteria lacking the Deformed expression plasmid (No Dfd) to the normal Deformed-containing extract (Dfd). The gel was intentionally overexposed to illustrate the degree of non-specific precipitation.
211 Targets of the Deformed homeotic gene

line to detect β-gal. As shown in Fig. 8, HZ-4.0 is capable of directing β-gal expression in cells of the maxillary lobe during late embryogenesis in a pattern very like the 1.28 enhancer trap except the enhancer trap element expression is nuclear (Fig. 8A) while HZ-4.0 is cytoplasmic (Fig. 8B). This pattern of staining in the maxillary lobe was observed in all four primary transformation lines examined. Though the 1.28 transcript does not accumulate at this late stage, the stability of β-gal from the enhancer trap element allows it to serve as a marker for cells previously expressing 1.28. Furthermore, though HZ-4.0 staining intensity is lower than that from the enhancer trap, making detection difficult at earlier stages, we are able to detect β-gal expression in the maxillary region of somewhat younger embryos than that shown, and this HZ-4.0 expression is in the same position as the 1.28 transcripts. Significantly, HZ-4.0 expression was not detected in all Deformed-expressing regions of the embryo (Fig. 8C). Maxillary expression was not seen when fragments lacking the Deformed binding region were tested nor in homozygous Deformed^{w21} embryos (data not shown). Taken together, the DNA binding and enhancer tester results strongly suggest that a direct interaction between Deformed and the 1.28 gene is required to activate this gene in cells of the maxillary lobe.

DISCUSSION

Homeotic genes specify segment identity by encoding transcription factors which select the appropriate battery of target genes necessary to determine the fate of a given segment (Morata and Struhl, 1990; Andrew and Scott, 1992). These target genes, either directly or indirectly, lead to the production of individual, segment-specific characteristics, e.g., sense organs, muscle pattern and specialized cuticle structures. In this manuscript we describe the identification of one such target gene, a gene activated by the Deformed homeoprotein within cells of the maxillary segment. Our evidence strongly supports the contention that activation of the 1.28 gene in the maxillary cells requires a direct interaction with Deformed. The 1.28 enhancer trap and the 1.28 transcript are expressed cell-for-cell with Deformed in the maxillary epidermis, and the Deformed protein is required for this expression. Further, an enhancer region, containing Deformed binding sites and capable of directing maxillary expression of a reporter gene, was found upstream of the 1.28 gene. The identification of a target gene that is activated by the homeoprotein is significant because most other target genes identified to date are repressed by the regulating homeoprotein (Gould et al., 1990; Wagner-Bernholz et al., 1991; Graba et al., 1992; Vachon et al., 1992).

Deformed is the first homeoprotein detected in the early embryo; it accumulates in a band of cells on either side of the cephalic furrow (Jack et al., 1988; Mahaffey et al., 1989). These cells will form the mandibular and maxillary segments. As germ band extension begins, Deformed protein expression expands anteriorly into the hypopharyngeal region. Once the germ band reaches full extension and begins to retract, segmentation becomes evident within the head region and Deformed becomes restricted to the posterior-lateral cells of the mandibular and maxillary lobes, while also accumulating in cells along the anterior edge of the dorsal ridge and in parasegment 1 of the ventral nerve cord. It is within these posterior-lateral maxillary cells, those maintaining Deformed expression, that we see the 1.28 gene expressed. Fate mapping studies have shown that
several larval mouthpart and sense-organ structures arise from the maxillary lobes (Jürgens et al., 1986). Whether any of these structures arise from the cells expressing 1.28 is unknown, but it seems doubtful considering the final location of the cells that have expressed the 1.28 gene. These maxillary cells remain outside the embryo after head involution and they will remain so in the first instar larvae.

Although Deformed is required to activate the 1.28 gene in certain cells of the maxillary lobe, there are several indications that Deformed alone is not sufficient for this activation. First, the 1.28 gene is expressed in only a limited number of Deformed-positive cells. We did not detect 1.28 expression in cells of the mandibular lobe, the dorsal ridge or in parasegment 1 of the ventral nerve cord, even though these cells express Deformed. Second, ectopic Deformed expression, after appropriate heat shock of embryos containing the hsDfd construct, did not activate β-gal expression from the 1.28 enhancer trap element, not even within those cells along the ventral region of the embryo that persistently express Deformed for the remainder of embryogenesis (Kuziora and McGinnis, 1988). Therefore, it appears that the environment necessary for Deformed to activate the 1.28 gene is specific to the posterior-lateral maxillary cells. Possibly, Deformed is only one component of an activation signal, and these cells possess other factors which are required, in combination with or prior to Deformed, to activate the 1.28 gene. Certainly, other mechanisms are plausible. Whatever the mechanism, since the 4.0 kb SalI fragment directs expression in a 1.28 pattern and not in all Deformed expressing cells, the genetic signals required for this limited expression seem to be present in the DNA of this fragment. Whether it is significant that we found potential binding sites for other developmental transcription factors within this enhancer region remains to be determined.

After observing the distribution of 1.28 transcripts in the embryo, one might predict that this gene is regulated by other homeotic genes in addition to Deformed, specifically Abdominal B, Antennapedia, and labial. Those cells expressing 1.28 in the first thoracic and eighth abdominal segmental segments are within the region of the embryo giving rise to the anterior and posterior spiracles (Campos-Ortega and Hartenstein, 1985). The posterior spiracles are specified by the Abdominal B homeoprotein; Abdominal B is expressed in these cells (Celniker et al., 1989) and the spiracles are absent in strong Abdominal B mutants (Sanchez-Herrero et al., 1985). Similar observations on the formation of the anterior spiracles (Wakimoto and Kaufman, 1981; Levine et al., 1983) indicate that 1.28 may be regulated by Antennapedia. Perhaps the homeodomain protein empty spiracles (Dalton et al., 1989) is also involved. However, whether these cells will indeed form the spiracles is uncertain at this time. Likewise, a case could be made for regulation of 1.28 by labial since 1.28 and early labial accumulation are similar in the midgut (Diederich et al., 1989). Perhaps, further studies will identify enhancers at the 1.28 gene that respond to these other homeoproteins.

Finally, though the target gene we have identified encodes a protein of unknown function, lethal mutations associated with small deletions in the upstream portion of the 1.28 gene (obtained by remobilizing the enhancer trap element) indicate that the gene is required during embryogenesis or early larval life (data not shown). The lack of existing deletions of the 42B region (see note in Materials and Methods) has hindered our genetic analysis, however, a complete analysis is underway. Searches of available data bases have not uncovered any significant similarities between the 1.28 open reading frame and other genes or proteins; however, similar amino acid compositions have been noted in other proteins. The homeoproteins Deformed and labial (Regulski et al., 1987; Diederich et al., 1989, respectively) have domains rich in serine, threonine and proline, as does the 1.28 protein between amino acids 71-130. This is also true for other transcription factors (for example, CTF/NF-I, Mermod et al., 1989; OTF-2, Gerster et al., 1990; Ap-2, Williams and Tjian, 1991; as well as others found in similarity searches using the BLAST algorithm). However, domains of this composition are not limited to transcription factors. Further, the protein encoded by the Hairless gene has a similar abundance of these amino acids (Maier et al., 1992) though the Hairless protein is much larger. Experiments are underway to address the role of the 1.28 gene product. Whether or not similar proteins are found in other animals will be of interest, especially considering the conservation of the Deformed homeoprotein between taxa (Graham et al., 1988; McGinnis et al., 1990). In any case, it is likely that the searches for target genes will identify previously unknown classes of proteins with functions as interesting and conserved as the homeoproteins before them.

We would like to acknowledge the following people for sharing their time and reagents which helped make this work possible; J. Gindhart and T. Kaufman for the enhancer tester vector pHZ-white, W. McGinnis for the materials necessary for the Deformed binding studies, the hsDfd fly line and the modified whole embryo in situ protocol, M. Conkling for the rapid phage filter protocol and M. Seeger and N. Patel for the PCR in situ probe generation protocol. These reagents were instrumental to this work. We also thank E. McGuckin for help in generating the enhancer trap lines, T. C. Kaufman, W. McGinnis and S. Celniker for helpful discussions, and K. Matthews and the Bloomington, IN stock center for supplying various fly stocks. We would like to thank K. Tatchell and D. Miller for comments on early versions of the manuscript. Protein similarity search computation was performed, in part, at the NCBI using the BLAST network service.

This work was supported by Grant #DCB-9108892 from the National Science Foundation.

REFERENCES


(Accepted 5 February 1993)