Development 111, 407-424 (1991)
Printed in Great Britain © The Company of Biologists Limited 1991

The large upstream control region of the *Drosophila* homeotic gene *Ultrabithorax*

KENNETH D. IRVINE, STEPHEN L. HELFAND* and DAVID S. HOGNESS†

Department of Biochemistry, Beckman Center, Stanford University School of Medicine, Stanford, California 94305-5307, USA

Present addresses: *Department of Biostructure and Function, University of Connecticut Health Center, Farmington, Connecticut 06032 USA; †Department of Developmental Biology, Beckman Center, Stanford University School of Medicine, Stanford, California 94305-5427, USA

Summary

***Ultrabithorax* (*Ubx*) is a *Drosophila* homeotic gene that determines the segmental identities of parts of the thorax and abdomen. Appropriate *Ubx* transcription requires a long upstream control region (UCR) that is defined genetically by the *bithoraxoid* (*bxd*) and *postbithorax* (*pbx*) subfunction mutations. We have directly analyzed UCR functions by the examination of β-galactosidase expression in flies containing *Ubx–lacZ* fusion genes. 35 kb of UCR DNA confers upon β-galactosidase an expression pattern that closely parallels normal *Ubx* expression throughout development. In contrast, 22 kb of UCR DNA confers fewer features of normal *Ubx* expression, and with 5 kb of UCR DNA the expression pattern has no resemblance to *Ubx* expression except in the visceral mesoderm. We have also shown that *bxd* chromosome breakpoint mutants form a comparable 5′ deletion series in which the severity of the effect on *Ubx* expression correlates with the amount of upstream DNA remaining in the mutant. In *Ubx–lacZ* fusions containing 22 kb of UCR DNA, and in comparable *bxd* mutants, there is a persistent pair-rule pattern of metameric expression in early development, demonstrating that there are distinct mechanisms with different sequence requirements for the initial activation of *Ubx* in different metameres. The correction of this pair-rule pattern later in embryogenesis shows that there are also distinct mechanisms for the activation of *Ubx* at different times during development.**

Key words: *Drosophila* development, bithorax complex, transcriptional regulation, *cis*-regulatory elements.

Introduction

Early in the development of *Drosophila melanogaster*, a number of maternal and zygotic genes divide the embryo into metameres and provide positional information for the region-specific activation of the homeotic genes (reviewed in Akam, 1987; Scott and Carroll, 1987; Ingham, 1988). The homeotic genes specify the metameric identities of the cells in which they are expressed; loss or alteration of homeotic gene expression transforms metameric identity (reviewed in Duncan, 1987; Kaufman *et al.* 1990). Understanding how the metameres adopt distinct fates thus requires an understanding of how homeotic gene expression is regulated. We are investigating how the homeotic gene *Ultrabithorax* (*Ubx*) interprets the positional information provided by earlier acting genes and how appropriate *Ubx* expression patterns are maintained and modulated throughout development.

Ubx is a member of the bithorax complex (BX-C), a cluster of homeotic genes that specify the identities of abdominal and posterior thoracic segments (Lewis, 1978; reviewed in Duncan, 1987). Segments and parasegments are metameres each composed of an anterior (a) and a posterior (p) compartment (García-Bellido *et al.* 1973, 1979) but are out of phase; segments comprise an a+p unit while parasegments comprise a p+a unit (Martinez-Arias and Lawrence, 1985). In the epidermis, *Ubx* is primarily responsible for determining the identities of parasegments 5 and 6 (PS5 and 6) in both the larva and the adult (PS5=T2p+T3a and PS6=T3p+A1a, where T2 and T3 are the second and third thoracic segments, and A1 is the first abdominal segment). In *Ubx* mutants these parasegments are transformed toward PS4 (T1p+T2a). In addition, *Ubx* plays a minor role in determining metameric identities in the larval epidermis posterior to PS6 that is secondary to that played by the other genes of the BX-C (Lewis, 1978). *Ubx* functions in internal tissues have been less well characterized, but for the larval visceral mesoderm *Ubx* is required in PS7 (Bienz and Tremml, 1988), and for the larval somatic mesoderm *Ubx* is required in abdominal segments (Hooper, 1986). In the nervous system, the metameric requirements for *Ubx*

function closely parallel those in the epidermis (Teugels and Ghysen, 1985). The spatial expression pattern of *Ubx* protein (UBX) correlates well with the genetic requirements but has also revealed complexities of UBX expression not predicted by genetic analysis (White and Wilcox, 1984, 1985*a*; Beachy *et al.* 1985; Brower, 1987; Canal and Ferrús, 1987).

The *Ubx* gene has been cloned and its products analyzed (Bender *et al.* 1983; Hogness *et al.* 1985; O'Connor *et al.* 1988; Kornfeld *et al.* 1989). The 77 kb *Ubx* transcription unit is alternatively spliced to produce at least 5 different mRNAs which share common 5′ and 3′ exon sequences but differ in their use of three internal elements. These mRNAs are translated to produce a family of UBX proteins which differ in their tissue and temporal distributions (J. Lopez and D.S.H., unpublished data). Each UBX protein contains the same homeodomain, a DNA-binding motif conserved among a large group of known and putative transcriptional regulators (reviewed in Scott *et al.* 1989). UBX proteins have been purified and shown to be sequence-specific DNA-binding proteins (Beachy *et al.* 1988), and can regulate reporter gene expression in a tissue culture cell cotransfection assay (Krasnow *et al.* 1989) and in an *in vitro* transcription system (Johnson and Krasnow, 1990).

Ubx is a complex locus with four classes of subfunction mutations, *anterobithorax* (*abx*), *bithorax* (*bx*), *bithoraxoid* (*bxd*) and *postbithorax* (*pbx*), each of which cause a subset of *Ubx* mutant phenotypes and are not complemented by *Ubx* mutations (Lewis, 1978; reviewed in Duncan, 1987). It has been proposed that most or all of the functions of the locus are provided by UBX proteins and that the subfunction mutations are lesions in *Ubx* regulatory sequences (Beachy *et al.* 1985; Hogness *et al.* 1985; Peifer *et al.* 1987). This view is supported by the observation that *Ubx* expression is altered or eliminated in the cells affected by a number of these mutations. (Beachy *et al.* 1985; Cabrera *et al.* 1985; Hogness *et al.* 1985; White and Wilcox, 1985*b*; Botas *et al.* 1988; Little *et al.* 1990). The *abx* and *bx* mutations primarily affect the development of PS5 and map downstream of the transcription start site (Bender *et al.* 1983; Peifer and Bender, 1986); we call this the downstream control region or DCR (Fig. 1A). The *bxd* and *pbx* mutations primarily affect the development of PS6 and more posterior parasegments and map upstream of the transcription start site (Bender *et al.* 1983, 1985; Lipshitz *et al.* 1987); we call this the upstream control region or UCR (Fig. 1A). The UCR DNA is also transcribed, but these transcripts play little if any role in UCR function (Hogness *et al.* 1985; Lipshitz *et al.* 1987).

Many of the *bxd* mutations are caused by chromosome breakpoints and are of particular interest because they form an allelic series in which breakpoints close to the *Ubx* transcription start site cause strong transformations while breakpoints further away cause weaker transformations (Bender *et al.* 1985). Breakpoints that cause *bxd* phenotypes are located from 2 to approximately 40 kb upstream of the transcription start site (Fig. 1A). This, together with the alterations in UBX expression in *bxd* mutants and the genetic requirement for *bxd* in *cis* to *Ubx* (Lewis, 1955), suggest that the UCR is a transcriptional regulatory region of extraordinary length. To test this suggestion, we have fused portions of the UCR to a reporter gene, *E. coli* lacZ, which encodes the enzyme β-galactosidase, transformed these fusion genes into *Drosophila* and monitored the resulting expression patterns. In addition, the UBX protein distribution in a series of *bxd* breakpoints spanning the UCR has been examined. Our results provide further insights into the regulation of *Ubx* expression and show that UCR functions are effected through transcriptional regulation of *Ubx*.

Materials and methods

Drosophila *strains*

Wild-type Canton-S embryos were used for injection of P-element constructs marked with the Tn5 *neo* gene, and w^{1118} embryos were used for injection of constructs marked with the $white^{+}$ gene. Stocks were obtained from G. Rubin, W. Bender, E. Lewis and the Indiana stock center. Descriptions of strains used in this work can be found in Hazelrigg *et al.* (1984) for w^{1118}; Robertson *et al.* (1988) for Δ2-3(99B); Lewis, (1978), Bender *et al.* (1985) and Casanova *et al.* (1985) for BX-C mutants.

Vector and fusion gene constructions

Derivatives of cosPneo (Steller and Pirrotta, 1985) were used as transformation vectors. cosPneoβ-gal, made by insertion of the *Sma*I–*Eco*RI *lacZ*–SV40 poly(A) fragment from pC4β-gal (Thummel *et al.* 1988) into cosPneo, was generously provided by H. Lipshitz. This vector was modified by the sequential removal of *Xba*I and *Sal*I sites outside the polylinker by digestion, endfilling and religation; the resulting vector was named cPnβb. cPwβ was constructed by joining the 9.4 kb *Bgl*II–*Eco*RI fragment of cPnβc to the 6.7 kb *Bgl*II–*Eco*RI fragment of cosper. cPnβc was derived from cPnβb by insertion of a CGGTACCG *Kpn*I linker into an endfilled *Sal*I site in the polylinker. cosper was a gift of V. Pirrotta and is a derivative of cosPneo in which the Tn5 *neo* gene has been replaced by the *Drosophila white* gene (Pirrotta, 1988).

For the fusion of UCR sequences, the plasmids pEMBL18St and pEMBL19St were constructed from pEMBL18 and pEMBL19 (Dente *et al.* 1983), respectively, by insertion into the *Cla*I site of multiple CCTCGAGG *Xho*I linkers; adjacent linkers generate a *Stu*I site. The constructs were designed to preserve the genomic *Ubx* sequence with no nucleotides added or deleted in cloning. The 1.6 kb *Stu*I–*Hin*dIII fragment of $\lambda_c$dm2228 (−32.0 to −30.4 kb on the BX-C molecular map; Bender *et al.* 1983) and the 4.6 kb *Hin*dIII–*Xba*I fragment of $\lambda_c$dm2218 (−30.4 to −25.8 kb) were sequentially cloned into pEMBL18St to generate Pbxd6.2. The Pbxd6.2 insert was removed by digestion with *Xho*I, treatment with S1 nuclease and large fragment DNA polymerase I, followed by digestion with *Xba*I. cPnβbxd6.2 was formed by cloning this fragment into *Sma*I–*Xba*I cut cPnβb; this generated an open reading frame fusion with the first eight codons of *Ubx* and three polylinker encoded codons joined to the eighth codon of *lacZ*. The fusion junction was verified by DNA sequencing. The 6.2 kb *Bam*HI–*Xba*I UCR fragment of cPnβbbxd6.2 was cloned into cPwβ to generate cPwβbxd6.2.

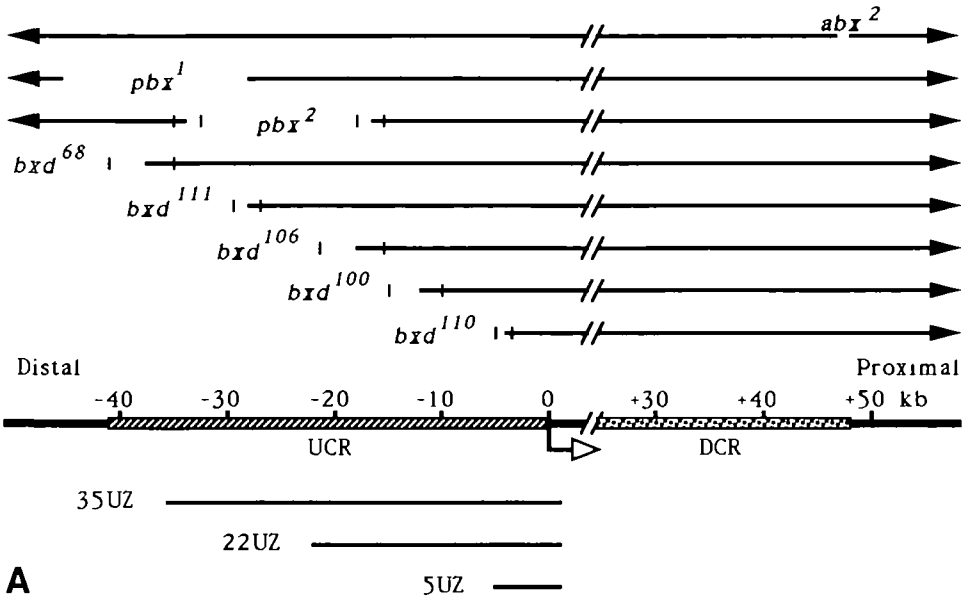
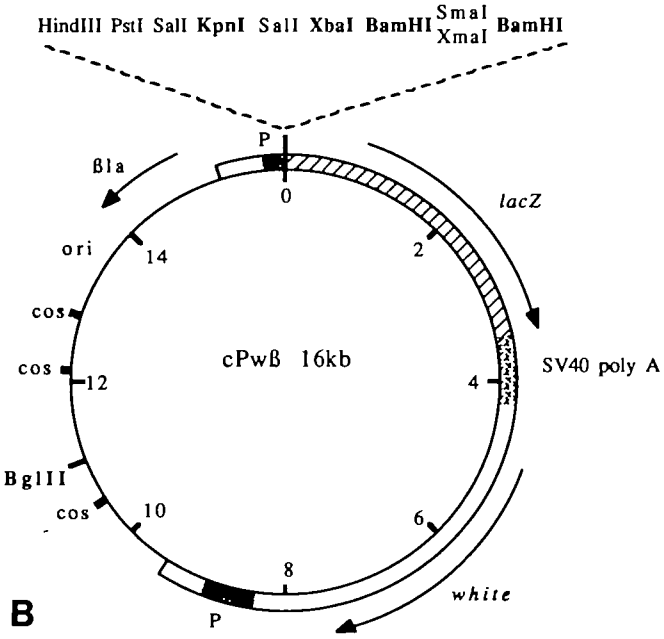


Fig. 1. Maps of *Ubx* and the transformation vector. (A) The coordinates relative to the *Ubx* transcription start site (open arrow; Kornfeld *et al.* 1989), chromosomal orientation, UCR and DCR are indicated. The UCR is currently defined as extending from the transcription start site to the bxd^{68} breakpoint (Bender *et al.* 1985), and the DCR is currently defined as extending from the most distal *bx* insertion to the most proximal end of *abx* deletions (Peifer and Bender, 1986). Above the map the extent of the DNA retained in various mutations is indicated by horizontal lines, with the vertical slashes indicating the uncertainty in the mapping of the DNA breakpoints. Below the map the extent of *Ubx* DNA in the *Ubx–lacZ* fusions is indicated. (B) The P element ends (P), sites for phage λ *in vitro* packaging (cos), bacterial origin of replication (ori), ampicillin resistance gene (βla), SV40 small t splice and polyadenylation signal (SV40 poly A), *E. coli* lac Z gene (*lacZ*), *Drosophila white* gene (*white*), and polylinker are indicated. Note that the P element ends are flanked by *white* DNA. Unique restriction sites are indicated in bold. In cPnβc, the *white* gene is replaced by Tn5 *neo*; in this vector the *Sma*I and *Sal*I sites become unique. Polylinker sequence: GGT CGA CGG TAC CGT CGA CTC TAG AGG ATC CCC GGG GAT CCC GTC, where GTC is the eighth codon of *lacZ*.



The DNA upstream of this 6.2 kb UCR fragment was fused by the sequential joining of three fragments, the 10.4 kb *Xba*I–*Bam*HI fragment of $\lambda_c$dm2218 (−25.8 to −15.4 kb), the 6.6 kb *Bam*HI–*Kpn*I fragment of $\lambda_c$dm2212 (−15.4 to −8.8 kb), and the 13.2 kb *Kpn*I–*Stu*I fragment of cosbxdA (K.D.I., unpublished data; −8.8 to +4.4 kb) in pEMBL19St. cPwβbxd36 and cPwβbxd23 were formed by the ligation of the resulting 30.2 kb *Xba*I–*Stu*I and 17.0 kb *Xba*I–*Kpn*I fragments, respectively, into cPwβbxd6.2.

Based on preliminary nucleotide sequence analysis of the UCR (D. Peattie, L. Prestidge, and D.S.H., unpublished data), the published restriction map (Bender *et al.* 1983) has been slightly revised. The 36.4 kb encompassed by cPwβbxd35 is expanded by 1.4 kb over what would have been determined from the published map. For the revised map, the *Ubx* transcription start site has been defined as −31.000 kb. For comparison, the *Eco*RI sites in Canton-S DNA (Bender *et al.* 1983) that are distal to this start site are placed at −27.49, −20.95, −15.96, −15.07, −13.40, −12.26, −11 (not in current version of sequence), −9.94, −6.52, −4.21, −1.40, and +3.88 kb in the revised map. In Fig. 1A, where the transcription start site is placed at 0 kb, the Fig. 1A coordinates are related to those for the revised map by $y=-x-31$, where y and x are the Fig. 1A and revised map coordinates, respectively.

P element transformation

Transformation was performed essentially as described by Spradling and Rubin (1982). For cPwβbxd5, 0.4 mg ml^{-1} of PIChs$\pi\Delta$2-3 helper DNA (Mullins *et al.* 1989) and 0.8 mg ml^{-1} of construct DNA were used. Because of problems with the high viscosity of the DNA solutions for the two larger constructs, 0.3 mg ml^{-1} of both helper and construct DNA were used. Inserts were mobilized by crossing to a marked Δ2-3(99B) chromosome. Putative chromosomal hops of the insert were identified by a change in eye color

intensity or pattern (caused by increase in copy number and position effects) or by a change in linkage group.

All lines were examined by genomic Southern to establish copy number, distinguish different insertion sites and check the integrity of the inserted DNA. Some hops in which the transposed construct was not intact were observed and discarded. In addition, the cytological locations of most inserts linked to the third chromosome were determined by *in situ* hybridization to polytene chromosomes as described by Zuker *et al.* (1985). The chromosomal locations of the constructs in the different lines are: 5UZ-1: X; 5UZ-2: 3; 5UZ-3: 3(98B); 5UZ-4: 3(89E); 5UZ-5: 2; 22UZ-1: 3(73F); 22UZ-2: 3(82E); 22UZ-3: X; 22UZ-4: 3(TM6bTb); 22UZ-5: 3(73D); 35UZ-1: 3(78A); 35UZ-2: 3(92D-E, TM3Sb); 35UZ-3: 2(34D). Line 5UZ-1 is marked with Tn5 *neo*, all other lines are marked with w^+. Although the insert in line 5UZ-4 is within 89E, which includes the cytological location of the BX-C (89E1–4; Lewis, 1978), it is not within the 35 kb region from −31 to +4 on the BX-C molecular map, has no bithorax phenotype and the expression pattern is not distinguishable from that of other 5UZ lines.

Determination of expression patterns

Spatial patterns of expression were determined by β-galactosidase activity stains, immunohistochemistry and immunofluorescence. *bxd* and *abx* mutant embryos were examined as homozygotes and as hemizygotes over Df Ubx^{109}, and were identified by their altered patterns of expression. The positions of β-galactosidase and UBX expression throughout development were determined by double-labelling experiments using antibodies directed against *en* protein.

β-galactosidase activity stains were performed as described in Hiromi *et al.* (1985), except that staining was performed in a pH 7 buffer at 28°C. After staining, embryos were rinsed in 70% and 100% ethanol, mounted in a 1:1 mixture of methyl salicylate and Canada Balsam, and photographed on a Zeiss photomicroscope with bright-field illumination or a Zeiss axiophot microscope with DIC optics.

For immunohistochemical stains embryos were dechorionated in bleach, fixed 20 min in heptane saturated with 4% paraformaldehyde, and devitellinized in methanol:heptane essentially as described by Mitchison and Sedat (1983). Embryos were often stored in methanol at −20°C prior to staining, and were stained by a horseradish peroxidase detection method (MacDonald and Struhl, 1986) with Ni/Co enhancement as described by Lawrence *et al.* (1987). This was the most sensitive method of detecting expression. Embryos were rinsed in BST [BSS (40 mM NaCl, 55 mM KCl, 10 mM $MgSO_4$, 1 mM$CaCl_2$, 10 mM tricine, 20 mM glucose, 50 mM sucrose, 0.2% BSA, 0.02% azide, pH 7)+0.1% Triton X-100], then washed 30 min each in BST and BSN (BSS+ 5% fetal calf serum+0.2% saponin). Incubation with primary antibody was in BSN for 2 h at room temperature or overnight at 4°C. Primary antibodies were diluted 1:2 to 1:5 for FP3.38 (anti-UBX *et al.* 1984), 1:4 to 1:10 for J3.1 (anti-UBX, J. Lopez and D.S.H., in prep.) 1:400 to 1:500 for 4D9 ascites (anti-*en*, Patel *et al.* 1989), and 1:800 to 1:2000 for rabbit anti-β-galactosidase (Cappel, preabsorbed against untransformed embryos). Both anti-UBX antibodies were monoclonal with epitopes common to all members of the UBX protein family. Embryos were then rinsed 4 times over a 1 h period in BST, 30 min in BSN (goat or rabbit serum, depending on the source of the secondary antibody, was usually substituted for calf serum), followed by incubation, as described for the primary antibodies, with biotinylated secondary antibodies (Vector, preabsorbed) in BSN at 1:400 to 1:500. Embryos were then rinsed in BST and three times (15 min each) in PTW [PBS (10 mM phosphate buffer, 130 mM NaCl, pH 7.5)+0.1% Tween-20], incubated in a 1:100 dilution in PTW of reagents A and B (Vectastain ABC kit) for 30 min, rinsed 3 times (5 min each) in PTW, incubated 5 min in 0.4 to 1 mg ml^{-1} diaminobenzidine tetrahydrochloride, and finally stained by the addition of 0.03% H_2O_2 and 0.03% each of $NiSO_4$ and $CoCl_2$. For double-labelling, the entire staining procedure was repeated except that $NiSO_4$ and $CoCl_2$ were omitted from the final step. After staining the embryos were rinsed in PTW, ethanol and mounted as described for β-galactosidase activity stains. Immunoflourescence was also used to detect expression; the protocol was similar up to the addition of the secondary antibody, when FITC, RITC or Texas-Red labelled secondary antibodies (Cappel) at 1:50 to 1:100 were added. Embryos were then often stained with bisbenzidine or DAPI to visualize the nuclei, and were mounted in 90% glycerol, 2% *n*-propyl gallate, PBS.

Dissected imaginal discs were stuck on poly-L-lysine coated slides, fixed in 4% paraformaldehyde, stained as embryos and mounted in permount after rinsing in ethanol and xylene.

Determination of β*-galactosidase specific activity*

Extracts of stage 11 embryos (5:20 to 7:20 h, timed collections) were made by homogenization in Z buffer (0.1 M phosphate buffer, 10 mM KCl, 1 mM $MgSO_4$, 50 mM β-mercaptoethanol, pH 7.0)+0.1% NP40. β-galactosidase activity was assayed as described in Miller (1972) and protein concentrations of extracts were determined by the method of Bradford (1976) using the reagent supplied by Bio-Rad. Measurements were made on at least three different lines of each construct and three separate extracts each assayed in duplicate for each line. The specific activities for the different lines, normalized to two copies per genome, were: 5UZ-2: 0.785; 5UZ-3: 0.818; 5UZ-4: 0.921; 5UZ-5: 0.652; 22UZ-1: 2.79; 22UZ-2: 4.07; 22UZ-3: 2.61; 35UZ-1: 9.91; 35UZ-2: 8.35; and 35UZ-3: 7.43.

Quantitation of CNS labelling

Stage 14–15 embryos were filleted to expose the nerve cord and stained by immunofluorescence using a protocol similar to that described for whole embryos. Quantitation of nuclei labelling was determined by direct microscopic counting of PS7–13 for bxd^{110}, bxd^{125}, bxd^{100}, PS12–13 for bxd^{113} and bxd^{106}, and PS13 for bxd^{111}, bxd^{98} and wild type. The increase in time required to count the greater numbers of nuclei in the other parasegments resulted in attenuation of the fluorescence signal and thus precluded direct microscopic counting. In order to evaluate these other parasegments, photographs of multiple specimens for each embryo type were projected and the numbers of nuclei labelling in each parasegment were counted blind three separate times. The reduced numbers of nuclei staining in the posterior parasegments of these *bxd* mutants tend to be in one focal plane allowing for an accurate assessment by use of photographs. This assumption was confirmed by the concurrence of the data obtained using both methods whenever possible.

Results

Construction and transformation of Ubx–lacZ *fusion genes*

The molecular genetic analysis of *bxd* and *pbx* mutants (Bender *et al.* 1983, 1985) suggested that a large amount of UCR DNA would be required for correct ex-

pression. Our goal in making the *Ubx-lacZ* fusions was first to construct fusions in which the UCR was of sufficient length to yield the wild-type expression pattern in order to define the complete regulatory region, and second to begin a dissection of this region into its component parts. In practice this approach was limited by the large size of this region and the technical size limits of the cloning and *Drosophila* transformation protocols that were employed. The three constructs described here, cPwβbxd5, cPwβbxd22 and cPwβbxd35, were made in the P element cosmid vector cPwβ (Fig. 1B). The *Ubx* DNA in these constructs is fused to the *lacZ* open reading frame (ORF) at codon 8 of the *Ubx* ORF and extends 6.2, 23.2 and 36.4 kb upstream from this fusion junction, respectively (Fig. 1A). Because the *Ubx* gene has a 960 bp 5′ untranslated leader (Kornfeld *et al.* 1989) these constructs extend, respectively, 5.2, 22.2 and 35.4 kb upstream from the *Ubx* transcription start site.

The constructs were transformed into the *Drosophila* genome by P-element-mediated transformation (Spradling and Rubin, 1982); transformants were selected by resistance to G418 (Steller and Pirrotta, 1985) or expression of w^+ (Pirrotta, 1988). Five cPwβbxd5 (or cPnβbxd5 – see Materials and methods) transformants were obtained from 85 fertile adults after coinjection with a helper plasmid encoding P element transposase, whereas single cPwβbxd22 and cPwβbxd35 transformants were obtained from 97 and 328 fertile adults, respectively, after similar treatment. Additional lines containing these last two constructs were generated to determine the influence of surrounding chromosomal sequences on expression patterns. Four and two additional lines containing cPwβbxd22 and cPwβbxd35, respectively, were isolated after mobilization of the inserted P element constructs with the transposase-supplying line Δ2–3(99B) (Robertson *et al.* 1988). The lines containing the fusion gene constructs are referred to collectively as 5UZ, 22UZ and 35UZ and individually as 5UZ-1 to −5, 22UZ-1 to −5 and 35UZ-1 to −3. The integrity of the inserts, their copy number and chromosomal positions were determined by genomic Southern analysis, by their linkage group and by *in situ* hybridization to polytene chromosomes (see Materials and methods).

Analysis of early embryonic expression patterns

The regulatory functions of UCR fragments fused to β-galactosidase were assayed by the examination of β-galactosidase expression patterns in transformed flies throughout embryonic development. Only those expression patterns that were constant among all lines of a given size construct are described; in addition, most lines had minor unique β-galactosidase expression which presumably reflected the influence of surrounding chromosomal sequences. The expression patterns were compared to UBX expression patterns in wild-type and in *bxd* mutants. Immunohistochemical, immunofluorescent and β-galactosidase activity stains were used to visualize expression patterns (see Materials and methods). *Drosophila* embryogenesis takes about 22 h at 25 °C and can be divided into stages by a number of morphological events that occur during development. The descriptions of Campos-Ortega and Hartenstein (1985) were used to stage embryos for the comparison of expression patterns; times in hours after fertilization have been inferred from the developmental stage.

Expression of Ubx *in early embryos – a temporal pattern*

The expresssion of UBX in wild-type embryos was analyzed to facilitate the description and analysis of β-galactosidase expression patterns in transformed lines. During most of embryogenesis, UBX is expressed in a complex spatial pattern in PS5–13, with PS6 having the greatest proportion of expressing cells (Fig. 2; White and Wilcox, 1984, 1985*a*; Akam and Martinez-Arias, 1985; Beachy *et al.* 1985; White and Lehmann, 1986). The parasegmental stripes of expression can also be described in terms of the temporal pattern of when expression first became detectable within different parasegments.

Ubx RNA has been detected ~2:30 h after fertilization (stage 5) in a broad distribution that peaks anteriorly in PS6 and rapidly decreases to low levels posteriorly (Akam and Martinez-Arias, 1985). Fig. 2A shows that UBX proteins were first detected as a single band in early stage 9 at ~3:45 h. Fig. 4A shows the position of this band relative to the multiple bands of the *engrailed* (*en*) protein which demarcate the p compartments and hence, by their anterior borders, define the parasegment boundaries (DiNardo *et al.* 1985; Ingham *et al.* 1985; Kornberg *et al.* 1985). Taken together, these two figures indicate that this initial UBX band occupies the posterior portion of PS6 but extends little, if at all, anteriorly into its p compartment. The anterior margin of *Ubx* RNA expression is also posterior to the PS6 anterior boundary (as defined by *fushi tarazu* expression, Akam and Martinez-Arias, 1985; Lawrence *et al.* 1987). Nor does UBX appear to extend appreciably into PS7, as determined either from the relative widths of the UBX band and PS6 (Figs 2A and 4A) or from direct examination of Fig. 4A. However, the weak expression of UBX at this stage would not allow the detection of low levels of UBX in parasegments posterior to PS6, as Akam and Martinez-Arias (1985) observed for early *Ubx* RNA expression. The lag of approximately one hour between initial UBX protein and *Ubx* RNA detection in PS6 has been attributed to the time required to transcribe the 77 kb *Ubx* transcription unit by Kornfeld *et al.* (1989), who proposed that this delay might function in the regulation of *Ubx* and other homeotic genes, which all have long transcription units.

Ubx protein was next detected in PS8, 10 and 12 to generate a pair-rule pattern (Fig. 2B), so called because of the relationship of this pattern to the pattern of function and expression of the pair-rule class of segmentation genes. A pair-rule pattern has also been observed for *Ubx* RNA (Akam and Martinez-Arias, 1985). The pair-rule pattern was transient, as protein

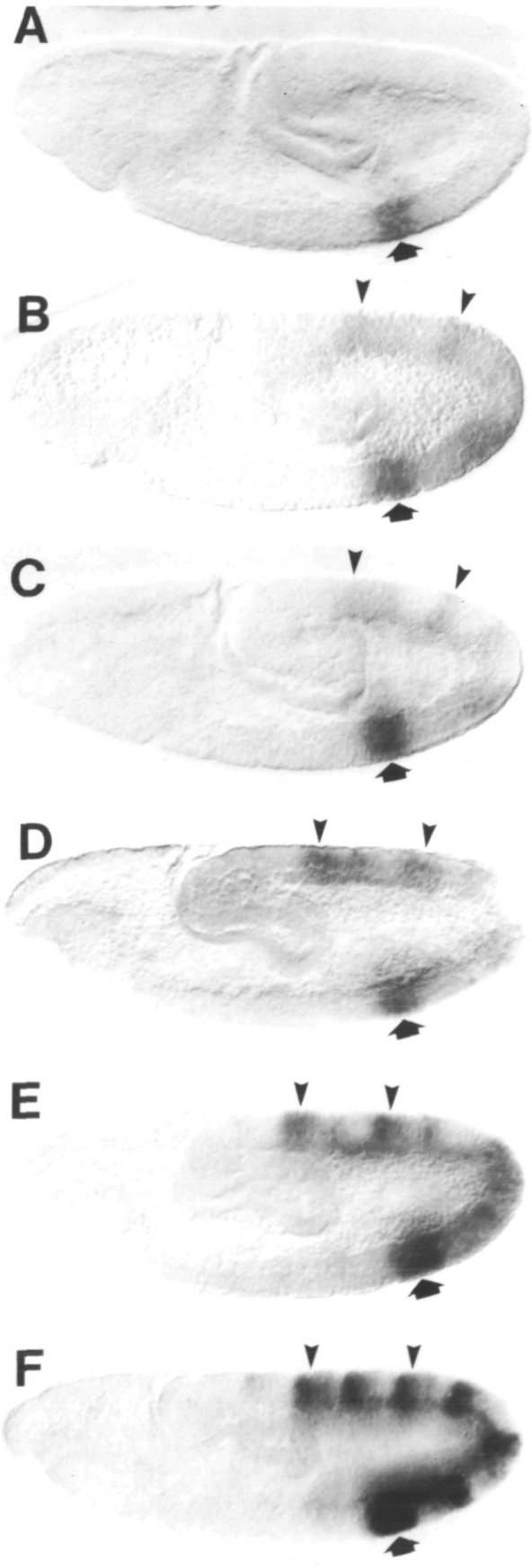


Fig. 2. Early embryonic expression of UBX in wild-type embryos, visualized by immunohistochemical staining. The embryos in this and all subsequent figures, unless otherwise indicated, are optical sections of whole-mount embryos with anterior to the left and ventral down. Staining in PS6 (arrow) and PS10 and 12 (arrowheads) is marked. The position of expression was based on double stains against *en* (Fig. 4 and data not shown). (A) Early stage 9 embryo with PS6 expression. The stage of earliest detectable UBX protein expression was based on the extent of germ band elongation and the morphology of the head, as it was difficult in these preparations to detect the transient mesodermal segmentation that distinguishes the end of stage 8 from the beginning of stage 9 (Campos-Ortega and Hartenstein, 1985). (B) Stage 9 embryo with pair-rule expression pattern in PS6, 8 10 and 12. (C) Late stage 9 embryo with additional expression detectable in PS9 and 11. (D) Stage 10 embryo with UBX expression in PS5–12; PS5 staining is not visible in this photograph as it was weak and out of the plane of focus. Expression in PS7, 8 and 9 was weaker than in PS10, 11 and 12 (see also Fig. 1E of White and Lehman, 1986). (E) Stage 11 embryo with expression in odd parasegments, which remained thinner and weaker than in even parasegments (compare PS10 and 12 with PS9 and 11). The faint smear of continuous expression in PS7 to 8 was reproducible. (F) Early stage 12 embryo. Equivalent expression was detected in each of PS 7–12. Age in hours after fertilization at 25 °C of different stages referred to in this work: 5, 2:10–2:50; 6, 2:50–3:00; 7, 3:00–3:10; 8, 3:10–3:40; 9, 3:40–4:20; 10, 4:20–5:20; 11, 5:20–7:20; 12, 7:20–9:20; 13, 9:20–10:20; 14, 10:20–11:20; 15, 11:20–13:00 (Campos-Ortega and Hartenstein, 1985).

was rapidly detected in PS9 and 11 (Fig. 2C) and then in PS5 and 7 (Fig. 2D). Further refinements in the expression pattern left PS6 as a strong broad band, with new expression detected in PS13 while PS7–12 formed a repeating ectodermal pattern in which UBX expression in each parasegment increased from weak to strong in progressing from the anterior to the posterior ends of the parasegment (Fig. 2E,F).

Expression of β-galactosidase in 35UZ parallels Ubx *expression*

In 35UZ, β-galactosidase was first detected at early stage 6 (~2:50 h after fertilization) in a stripe near the middle of the embryo (Fig. 3A). This expression was also localized relative to *en* expression; the sixth *en* stripe was separated by approximately one cell width of unstained cells from β-galactosidase (Fig. 4B). Thus β-galactosidase expression also initiated in the posterior half of PS6; it later expanded anterior to fill PS6 (Fig. 3B–D). The earlier detection of β-galactosidase *versus* UBX is consistent with the hypothesis that the size of the transcription unit significantly delays protein expression. If the difference were entirely due to the difference in sizes of the transcription units, 5 *versus* 77 kb, the rate of transcription of RNA polymerase II at 25 °C would be approximately 1.3 kb min^{-1}. This number could be skewed by differences in antibody sensitivity, protein or RNA stability, or post-transcriptional events, but nonetheless agrees well with an

estimate of 1.1 kb min^{-1} (±30%) obtained for an ecdysone-inducible gene (Thummel *et al.* 1990).

Despite this difference in expression delay times, the subsequent temporal and spatial patterns of β-galactosidase expression closely paralleled those for *Ubx* expression. β-galactosidase appeared in stage 7 as stripes in PS10 and 12, and in stage 8 expression appeared in PS8 (Figs 3B–D and 4C), to generate a PS6, 8 10, 12 pair-rule pattern. This pair-rule pattern was also transient; after PS8 expression, β-galactosidase was soon detected in PS9 and 11, then in PS5, PS7 and finally in PS13 by stage 10 (Fig. 3D–F and data not shown). β-galactosidase was also expressed, like UBX, in PS6–12 of the somatic mesoderm (Figs 2F and 3F).

This expression of β-galactosidase in a UBX pattern in 35UZ demonstrates that sequences sufficient for correct *Ubx* expression in early embryos are contained within the 35 kb of UCR DNA plus the 1 kb of 5′ leader DNA present in 35UZ. However, 35UZ lacks both upstream and downstream sequences known from genetic analysis to be required for proper *Ubx* function. To confirm that the β-galactosidase expression was consistent with the effects of mutations that remove *Ubx* sequences lacking in 35UZ, the expression of UBX protein in mutant embryos was examined. White and Wilcox (1985*b*) reported that the DCR mutants abx^2, abx^1 and bx^3 did not affect epidermal expression in late embryos. We examined early abx^2 embryos and could not distinguish the staining pattern from wild type, although in stage 13 embryos it did appear that PS5 epidermal expression was weaker (not shown). Thus, although regulation of UBX expression in PS5 has been attributed to the DCR, the analysis of UBX expression in DCR mutants is consistent with our observation that sequences sufficient for ectodermal PS5 expression in the early embryo are contained within the UCR. These PS5 regulatory sequences appear to be redundant, as DCR-*lacZ* fusions also generate expression within PS5 (Simon *et al.* 1990), and *bxd* mutants do not cause homeotic transformations in PS5. The bxd^{68} mutation, which has a similar amount of upstream DNA to 35UZ (Fig. 1A), was also not distinguishable from wild type in its staining patterns (not shown). This mutant has only a weak *bxd* phenotype, which could be caused by a slight decrease in the level of UBX expression which immunohistochemical staining was not sensitive enough to detect.

22UZ has altered spatial and temporal expression patterns

Having determined that 35UZ contains sequences sufficient to generate a *Ubx* expression pattern, the effect of removal of UCR sequences on the pattern of β-galactosidase expression was examined. In 22UZ, β-galactosidase was first detected at stage 6 in three stripes of approximately equal intensity in PS6, 10 and 12, and one very weak stripe in PS8, which became clearly visible in stage 7 (Figs 3G, H and 4D). This pair-rule pattern differed from that of 35UZ not only in the timing of its appearance but in the shape and relative intensity of the stripes of expression (Fig. 3J; cf. to Fig. 3D). Even more dramatically, while the 35UZ pair-rule pattern was transient, lasting only minutes, the 22UZ pair-rule pattern persisted until stage 11, lasting hours, with no new stripes of expression observed except for PS5 expression in stage 10 (Fig. 3J–L). Thus, some sequences required for initial activation of expression in PS7, 9 and 11 are located between 22 and 35 kb upstream of the transcription start site.

A transition in 22UZ expression occurred in stage 11 with the addition of ectodermal expression in PS7, 9, 11 and 13 so that in stage 12, β-galactosidase became expressed at equal levels in even and odd parasegments from PS7 to PS12 (Figs 5A and 6B). Thus some sequences sufficient for later expression in odd parasegments are located within 22 kb of the transcription start site. However, this later 22UZ expression pattern differed from that of 35UZ and UBX in that some differences in the intraparasegmental modulation of expression were observed (Fig. 5A; cf. to Figs 2F and 3F). Significantly, 22UZ did not have the strong broad band of expression in PS6 observed for 35UZ, the PS6 expression remaining more like that of PS7–12 (Fig. 5A; cf. to Fig. 3F). In addition, during stage 12, 22UZ began to express β-galactosidase in some ectodermal cells of PS4 and 5 in which 35UZ and *Ubx* are not expressed (Figs 5A and 8G), indicating the absence in 22UZ of sequences required to repress this ectopic anterior expression.

5UZ – a novel pattern

The further removal of UCR sequences in 5UZ abolished any relationship between β-galactosidase and UBX expression in the ectoderm. The pattern of β-galactosidase expression in 5UZ was similar throughout embryogenesis to that previously described by Bienz *et al.* (1988) for a *Ubx–lacZ* fusion gene with 3 kb of UCR DNA (Fig. 4E and data not shown). A prominent feature of the expression pattern generated by these short *Ubx–lacZ* fusion genes is the stripes of β-galactosidase detected in the lateral ectoderm of posterior compartments (Fig. 4E). The posterior stripes generated by 5UZ were first observed at stage 10 in PS5, 7, 9, 11 and 13, and soon became detectable in PS2–14, with those in PS7–11 being noticeably stronger. Any expression stripes anterior and posterior to this would have been obscured by the ectopic head and tail staining that is common to all constructs and is described in the last section of the results.

Bienz *et al.* (1988) showed that this striped pattern is dependent upon *Ubx* sequences proximal to the transcription start site, suggesting that the 5UZ pattern may reveal an aspect of the normal regulation of *Ubx*. For example, *Ubx* might be regulated by a transcription factor distributed in the 5UZ pattern. 5UZ would contain sequences sufficient to interact with this factor but not with other factors that normally modulate its influence. Alternatively, the pattern might reflect the response to such factors of a cryptic enhancer in the vector DNA that acts in conjunction with promoter-proximal *Ubx* sequences – a response that is masked by

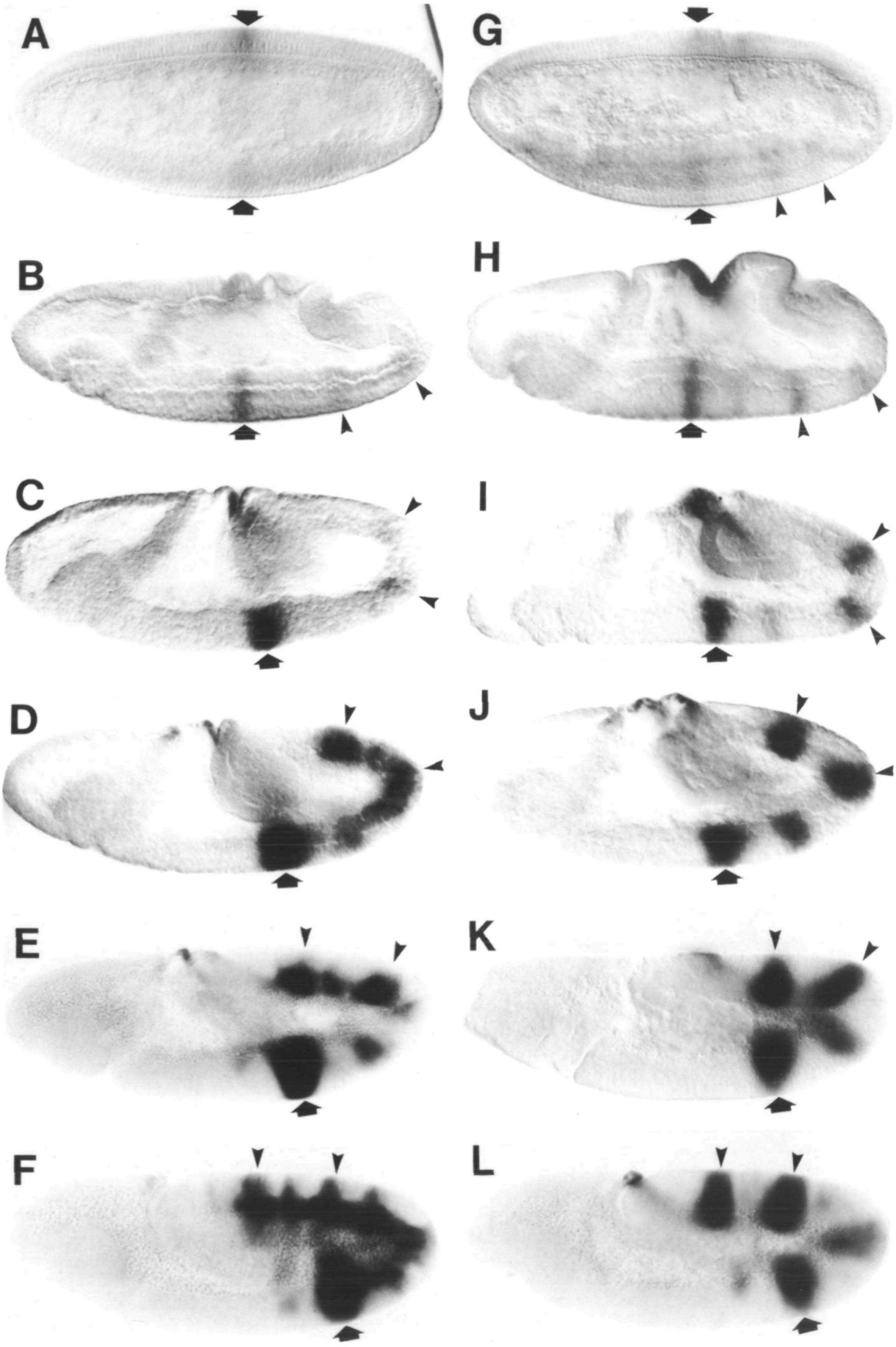

A
B
C
D
E
F
G
H
I
J
K
L

Fig. 3. Early embryonic expression of β-galactosidase in 35UZ (A–F) and 22UZ (G–L). A–D and G–J are immunohistochemical stains. E,F,K and L are β-galactosidase activity stains that have been stained more weakly than earlier embryos to allow comparison of the spatial patterns. Staining in PS6 (arrow) and PS10 and 12 (arrowheads) is marked. (A) Early stage 6 embryo; a single stripe of expression in PS6 was visible. This expression was stronger dorsally than ventrally. These dorsalmost cells will form part of the amnioserosa in which UBX and β-galactosidase were expressed later in development. (B) Stage 7 embryo with expression detectable in PS6, 10 and 12. (C) Early stage 8 embryo in which PS6 expression has increased and broadened. (D) Late stage 8 embryo with pair-rule pattern of expression in PS6, 8 10 and 12; a faint smear was visible in the lateral ectoderm throughout PS8 to 12 and there were weaker peaks of expression in PS 9 and 11. (E) Stage 9 embryo, expression was now also detected in PS5. (F) Stage 11 embryo with the expression stripes in PS7–12 almost equal. (G) Stage 6 embryo in which expression was detectable in PS6, 10 and 12, and was visible in PS8 in this embryo but at too low a level to reproduce in this photograph. (H) Stage 7 embryo in which expression is visible in PS8. (I) Early stage 8 embryo. (J) Late stage 8 embryo. Note the persistence of the strict pair-rule pattern and the general differences with 35UZ in the shape of the expression stripes (compare with Fig. 3D). (K) Stage 9 embryo. (L) Stage 11 embryo with expression just becoming visible in odd parasegments.

other UCR squences. Such a mechanism would be consistent with the observation that 5UZ-like patterns frequently result from the random insertion of *lacZ* genes in enhancer trap experiments (Bellen *et al.* 1989). Whatever the mechanism, the essential point is that 5UZ exhibits an aberrant ectodermal expression pattern and that this pattern can be converted to a *Ubx*-like pattern by addition of upstream UCR sequences.

Ectodermal expression of UBX in bxd *mutant embryos*

The effect of removing UCR sequences from *Ubx–lacZ* fusion genes was compared with the effect of removing UCR sequences from the chromosomal *Ubx* gene by the examination of *bxd* mutants (Fig. 1A). All *bxd* mutants other than bxd^{68} were distinguishable from wild type by altered patterns and decreased levels of UBX expression in PS6–13, although PS5 appeared unaffected. The decreased expression made it difficult to examine expression patterns in most mutants prior to stage 11, and in bxd^{110} prior to stage 12. At stage 11, bxd^{106} and bxd^{100} had a pair-rule expression pattern in PS6–12 similar to that seen in 22UZ (Fig. 5B; cf. to Fig. 3L). Furthermore, the remaining stripes in PS6, 8 10 and 12 were altered, being broad in the lateral ectoderm but sharply narrow in the ventral ectoderm. By contrast, in bxd^{111} the late pair-rule pattern was not observed, and although the level of expression was decreased relative to wild type, only in PS6 was the pattern of the parasegmental stripe noticeably altered (Fig. 5C). This implies that the 6.0±1.4 kb region between the upstream end of 22UZ and the bxd^{111} breakpoint interacts with some factor(s) required for early expression in PS7, 9 and 11. The late pair-rule pattern was not, however, observed in pbx^2 mutants (not shown), in which ~17 kb of UCR DNA is deleted, including this region. The pair-rule pattern in bxd^{106} and bxd^{100} was lost during stage 12 (Fig. 5D), slightly later (consistent with the transcriptional delay) than when the transition from the pair-rule pattern in 22UZ occurred. The detectable level of UBX expression at stage 12 appeared to correlate with the amount of upstream DNA retained in the mutant; that is, $bxd^{68} > bxd^{111} > bxd^{106} > bxd^{100} > bxd^{110}$ (Fig. 5D, E and data not shown).

The *Ubx* expression pattern in *bxd* mutants was generally consistent with β-galactosidase expression in *Ubx–lacZ* fusions with similar amounts of upstream DNA. However, the ectopic expression in PS4 and 5 seen in 22UZ was not observed in *bxd* mutants (Fig. 5D), and while bxd^{110} and 5UZ have similar amounts of upstream DNA, the expression patterns were completely different, as UBX expression in bxd^{110} was qualitatively similar to other *bxd* mutants (Fig. 5E). These observations again suggest the existence of some functional redundancy between the UCR and the DCR, as the abnormalities seen in the smaller fusion genes can be corrected either by the addition of DNA upstream (as in 35UZ) or downstream (as in a *bxd* mutant). Consistent with this suggestion, Simon *et al.* (1990) have observed that DNA fragments from the DCR generate expression patterns with similarities to *Ubx* expression when fused to a reporter gene.

Levels of β-galactosidase expression correlate with the amount of UCR DNA in Ubx-lacZ *fusions*

There were significant differences in the levels of expression in 5UZ, 22UZ and 35UZ, as judged by the time required to achieve approximately equivalent levels of staining in β-galactosidase activity stains. These differences in expression were directly compared by fixing and staining together 5UZ, 22UZ and 35UZ embryos (Fig. 6). Additionally, the differences were quantified by assaying the β-galactosidase specific activity in whole-embryo extracts. The mean specific activities in stage 11 embryos of 35UZ, 22UZ and 5UZ lines were 8.6, 3.1 and 0.79 units mg^{-1}, respectively. Although these values must be considered together with the differences in pattern among the lines, both the activity stains and the specific activities in embryo extracts indicate that the UCR controls the level as well as the pattern of expression.

Analysis of mid-embryonic expression patterns

Correct spatial patterning in the visceral mesoderm requires less UCR DNA than in other tissues

In the visceral mesoderm (the musculature that surrounds the gut), expression was detected in all lines in PS7, coincident with *Ubx* expression in this tissue (Fig. 7A–C and data not shown; Bienz *et al.* 1988). Consistent with this, bxd^{68}, bxd^{111}, bxd^{106}, bxd^{100} and bxd^{110} mutants were examined and found to be expressed in the visceral mesoderm in PS7 (Fig. 7D and data not shown), although the expression was very

weak in bxd^{110} mutants. The contrast between the correct expression pattern of 5UZ in the visceral mesoderm and the absence of a *Ubx*-like pattern in the ectoderm demonstrates that the sequence requirements for correct spatial patterning of *Ubx* in different tissues are distinct; correct epidermal expression patterns required greater than 22 kb of UCR DNA while the correct visceral mesoderm pattern required less than 5 kb. The regulation of visceral mesoderm also appears to be simpler than the regulation of ectodermal expression; each of the *Ubx–lacZ* genes had the same pattern of visceral mesoderm expression but distinct patterns of ectodermal expression. The sequence requirements for this visceral mesoderm expression have been further analyzed by J. Müller *et al.* (1989).

Regulation of central nervous system expression

In the central nervous system (CNS), UBX is expressed in a complex pattern similar to that of earlier epidermal expression (Fig. 8A; White and Wilcox, 1984, 1985*a*; Akam and Martinez-Arias, 1985; Beachy *et al.* 1985; Canal and Ferrús, 1987). There is strong expression in PS6 and a repeated parasegmental pattern of expression in each of PS7–12 in which the level of expression decreases slightly in the more posterior parasegments. PS5 and 13 have weaker and unique patterns of expression. Within any parasegment there is heterogeneity in the level of expression from cell to cell. The description of this expression pattern can be simplified by the observation that the few nuclei in PS13 that express UBX are in homologous positions to those of PS7–12 that show strongest expression, and thus the apparent difference in pattern could be a quantitative effect. Furthermore, to a first approximation the pattern in PS6 appears to be a composite of the PS5 and PS7 patterns; that is, the UBX-expressing nuclei in PS6 appear to occupy positions homologous to those in both PS5 and PS7. Two basic patterns thus emerge: a 'thoracic pattern' in PS5 and 6 (T2 and T3) and an 'abdominal pattern' in PS6–13 (A1–A8; Fig. 8H). There are also some additional nuclei in PS6 that express UBX, and the level of nuclear expression in PS6 is greater than in PS5 or 7. Finally, a cluster of ventral midline nuclei at the anterior parasegmental border of PS4, 5 and 6 express UBX.

In all *bxd* mutants except bxd^{68}, the expression of the abdominal pattern was noticeably decreased in PS6 to 13 (Fig. 8B–D; Beachy *et al.* 1985; White and Wilcox, 1985*b*). The nuclei that still expressed UBX in these mutants appeared to be those nuclei that expressed UBX at the strongest levels in wild type, suggesting that the change in pattern could result from a quantitative effect (Fig. 8A–D and data not shown). The amount of detectable expression in different *bxd* mutants correlated with the amount of remaining upstream DNA (Figs 1A and 8B–D). This effect was quantified by counting the number of nuclei that expressed detectable levels of UBX in PS7–13 in wild type and in a series of 7 *bxd* mutants that span the UCR (Table 1). The clear distinctions among these mutants in the level of UBX expression indicate that there are many *cis*-regulatory

Table 1. *Number of anti-UBX-labelled nuclei in PS7-13 wild-type and* bxd *mutant midembryonic nerve cords*

Genotype	PS7	PS8	PS9	PS10	PS11	PS12	PS13
bxd^{110}	6.6	6.2	5.6	4.6	3.6	0.6	0
n=12	(2.4)	(2.0)	(1.4)	(1.6)	(1.8)	(1.0)	–
bxd^{125}	17.8	15.2	12.4	11.6	9.8	2.6	0
n=12	(7.4)	(5.3)	(3.4)	(2.8)	(2.6)	(0.6)	–
bxd^{100}	23.8	21.0	18.2	16.6	13.8	4.8	0
n=17	(4.8)	(4.6)	(5.2)	(5.0)	(3.8)	(1.6)	–
bxd^{113}	38.8	34.8	30.8	29.2	26.0	15.6	0.2
n=14	(4.4)	(4.4)	(3.2)	(3 0)	(3.2)	(2.0)	(0.6)
bxd^{106}	48.0	42.0	37.6	34.2	27.2	21.4	0.2
n=7	(4.1)	(3.2)	(2.0)	(3.0)	(5.0)	(2.0)	(0.6)
bxd^{111}	>300	>300	>300	>300	>300	>300	11.0
n=4							(1.4)
bxd^{68}	>300	>300	>300	>300	>300	>300	9.5
n=4							(1.9)
wild-type	>300	>300	>300	>300	>300	>300	11.0
n=7							(2.8)

Genotype refers to the homozygous genotype of the embryo; n equals the number of embryos counted for each genotype. The data are presented as the arithmetic mean of the number of nuclei counted in each parasegment, with the standard deviation indicated in parenthesis under each mean. The number of nuclei labelled in PS7-12 for bxd^{111}, bxd^{68} and wild type were too high to count precisely and therefore an estimate of greater than 300 is used. The bxd^{113} and bxd^{125} breakpoints, which are not shown in Fig. 1A, have been mapped to regions of the UCR overlapping the locations of the bxd^{106} and bxd^{100} mutations, respectively.

sequences spread across the UCR that affect the level of *Ubx* expression in the CNS. Interestingly, the deletions pbx^1 (White and Wilcox, 1985*b*) and pbx^2 (data not shown) did not have detectable effects on CNS expression, suggesting that more distal sequences can substitute for the deleted sequences.

The pattern in PS5 appeared to be unaffected by *bxd* mutations. Those cells in PS6 homologous to the PS5-expressing cells (the thoracic pattern) also appeared to be unaffected (Fig. 8B–D); as a result, in strong *bxd* mutants the pattern in PS6 largely recapitulates that of PS5, indicating that the abdominal and thoracic patterns of PS6 are under separate regulatory control. It has been reported that *bxd* mutations led to increased UBX expression in PS5 (Beachy *et al.* 1985; White and Wilcox, 1985*b*). While it sometimes appeared that PS5 expression was increased, upon closer examination we concluded that the apparent increase was only relative to the decreased PS6 expression; a similar conclusion has been reached by W. Bender (personal communication). In *abx* mutants, UBX expression was eliminated in most of PS5 (Fig. 8E; White and Wilcox, 1985*b*), although the homologous cells in PS6 appeared unaffected. None of the mutants that have been examined eliminated expression in the ventral midline nuclei of PS4, 5 and 6 (Fig. 8B–E; White and Wilcox, 1985*b*).

In 35UZ, β-galactosidase expression in the CNS was similar to that of UBX in PS6–12 (Fig. 8F), although a precise comparison of the patterns was difficult because β-galactosidase is cytoplasmic while UBX is nuclear. Strikingly, β-galactosidase was not expressed in PS5,

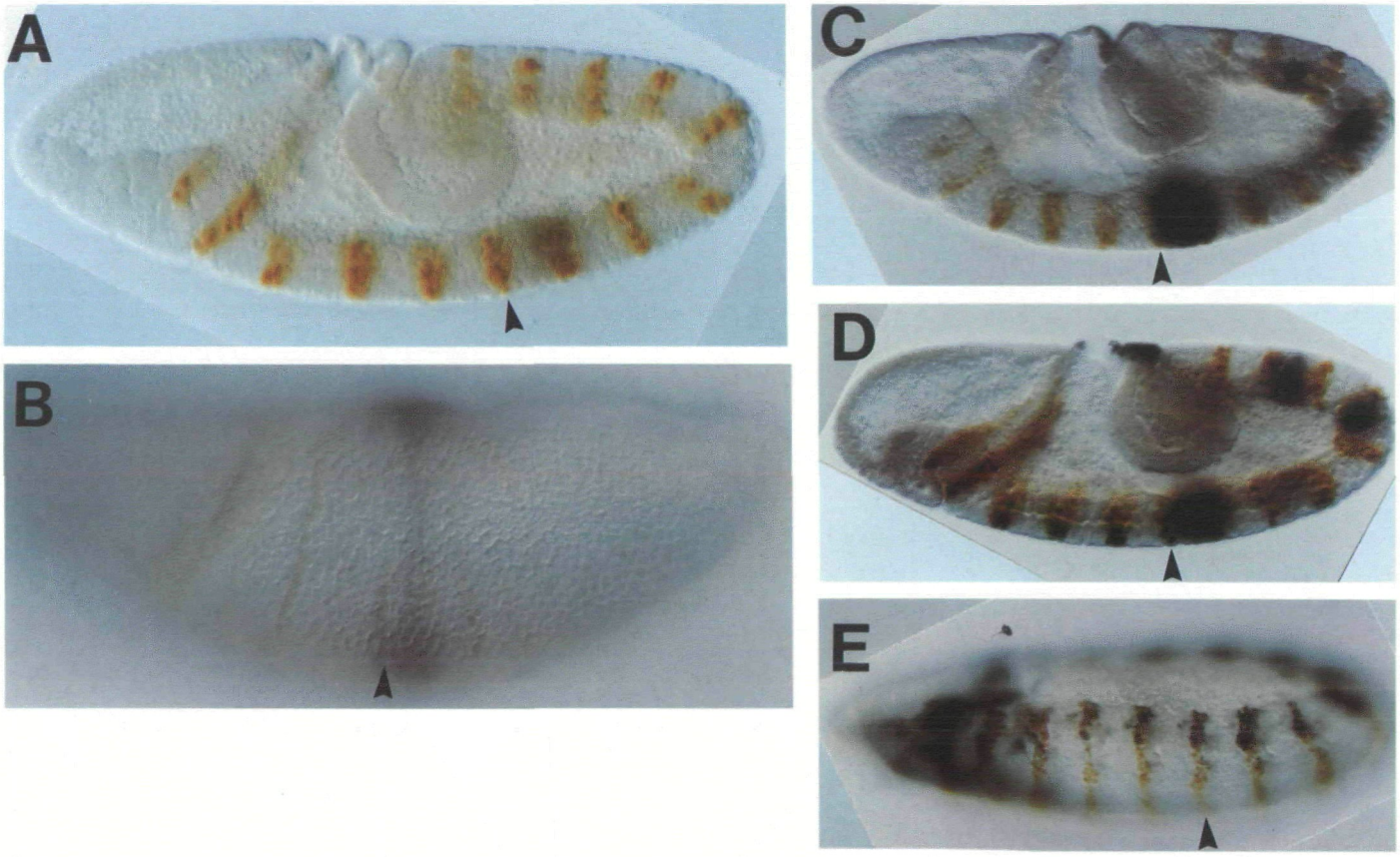


Fig. 4. Double labelling with anti-*en* and anti-UBX (A) or anti-β-galactosidase (B–E) antibodies. *en* protein was detected as orange while β-galactosidase and UBX were grey. The PS6 engrailed stripe, whose anterior margin marks the anterior boundary of PS6, is marked by the arrowhead. (A) UBX expression in stage 9. UBX was weak or nonexistent immediately posterior to the sixth *en* stripe, but increased posterior to that. The low level of expression prevents determination of the extent, if any, to which expression extended into PS7. (B) 35UZ at stage 7, focused on the ventrolateral surface. *en* expression stripes are one cell wide and have a pair-rule modulation at this stage, with expression barely visible in odd parasegments. β-galactosidase expression in the posterior half of PS6 was visible. (C) 35UZ at stage 8 showing a PS6, 8 10, 12 pair-rule pattern, with weaker expression detected in odd parasegments. (D) 22UZ at stage 8 with strict pair-rule pattern. (E) 5UZ at early stage 12, focused on the lateral surface, showing restriction of β-galactosidase stripes to the dorsolateral epidermis where they overlapped with *en*.

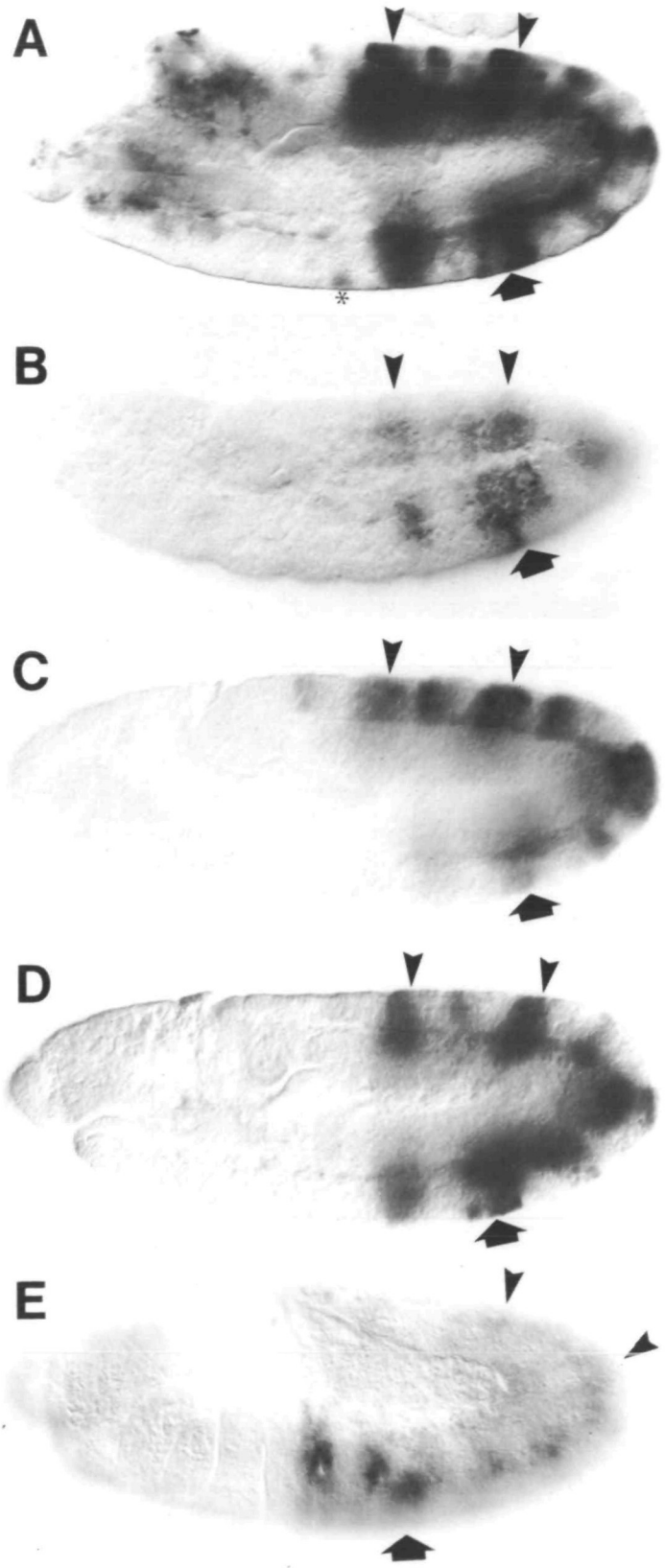


Fig. 5. UBX expression in *bxd* mutant embryos and comparison with β-galactosidase expression in 22UZ. Expression was visualized by immunohistochemical stains; PS6 (arrow) and PS10 and 12 (arrowheads) staining is marked. (A) β-galactosidase expression in 22UZ in early stage 12, showing addition of expression in odd parasegments. Note the ectopic expression in PS4 (asterisk). Comparisons of PS5 expression indicated ectopic expression there as well (see also Fig. 8G). (B) *Ubx* expression in bxd^{106} embryo at stage 11 with pair-rule pattern, focused on lateral surface. PS5 expression is also detected. (C) *Ubx* expression in bxd^{111} embryo at stage 11; the late pair-rule pattern was not observed. Most PS6 staining is out of the plane of focus, the pattern in PS6 was similar to that of bxd^{106}. (D) *Ubx* expression in bxd^{106} embryo at stage 12 with additional expression detected in PS7, 9 and 11. The apparent increase in PS5 expression relative to wild type (Fig. 2F) is a consequence of the increased staining necessary to compensate for the lower level of expression in PS6–13. (E) *Ubx* expression in bxd^{110} embryo at stage 12, focused on the lateral surface.

consistent with the effect of *abx* mutants and the absence of the entire DCR from the fusion gene. The ventral midline cells of PS4, 5 and 6 also did not express β-galactosidase. Taken together with the presence of UBX protein in these cells in *bxd*, *abx* and *bx* mutants, this result indicates that expression in these cells requires regulatory sequences downstream of the transcription start site which are not inactivated by existing mutations.

The band of β-galactosidase expression in PS6 appeared to be slightly thinner than UBX expression in PS6 (cf. Fig. 8F to E). To compare the expression patterns more accurately, the expression of β-galactosidase in 35UZ and of UBX in *abx* mutants were compared relative to *en*. This confirmed that some UBX- and *en*-expressing cells of PS6, in addition to the ventral midline cells, did not express β-galactosidase (data not shown). To a first approximation, the difference in PS6 expression between β-galactosidase in 35UZ and UBX in *abx* could be described as a loss of the thoracic pattern. However, the resolution of β-galactosidase expression in the CNS was not sufficient to allow a rigorous examination of this conclusion.

In 22UZ, fewer cells in PS6–12 in the CNS expressed β-galactosidase (Fig. 8G; cf. to Fig. 8F), similar to the effects of *bxd* mutants on UBX expression. The reduction was greatest in PS6, which had an expression pattern like that of PS7–12. Some cells in PS4 and 5 did express β-galactosidase, consistent with the ectopic epidermal expression detected in those parasegments. β-galactosidase was not detected in the CNS in 5UZ.

Expression of Ubx–lacZ *fusions in imaginal discs*

During larval life, the precursors of the adult structures in *Drosophila* are maintained as separate imaginal tissues. UBX expression in imaginal cells is expected to reflect the requirements for UBX proteins during the development of adult structures. UBX is expressed strongly in the T3 (haltere and third leg) imaginal discs and more weakly in certain posterior parts of the T2 (wing and second leg) imaginal discs (White and Wilcox, 1984, 1985*a*; Beachy *et al.* 1985; Brower, 1987).

In 35UZ, β-galactosidase was expressed strongly in the posterior of the T3 imaginal discs, that is, the PS6

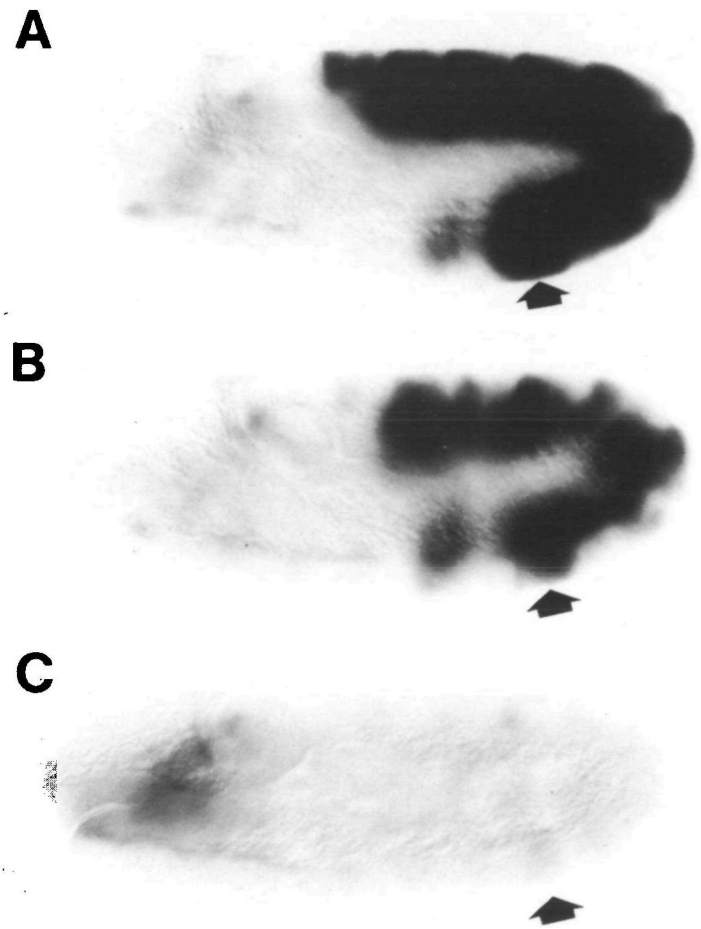


Fig. 6. Comparison among 5UZ, 22UZ and 35UZ. The stage 11 embryos in this figure were fixed and stained together in the same test tube for three hours with a β-galactosidase activity assay. The assay is not necessarily linear over this time so this is a qualitative rather than quantitative comparison. The arrow marks PS6. (A) 35UZ (B) 22UZ (C) 5UZ – the epidermal stripes are barely visible as they are very weak and out of the plane of focus in this figure. Note that while the level of expression in the germ band varies considerably among the three constructs, the head expression was similar in all three, although it was reproducibly greatest in 5UZ.

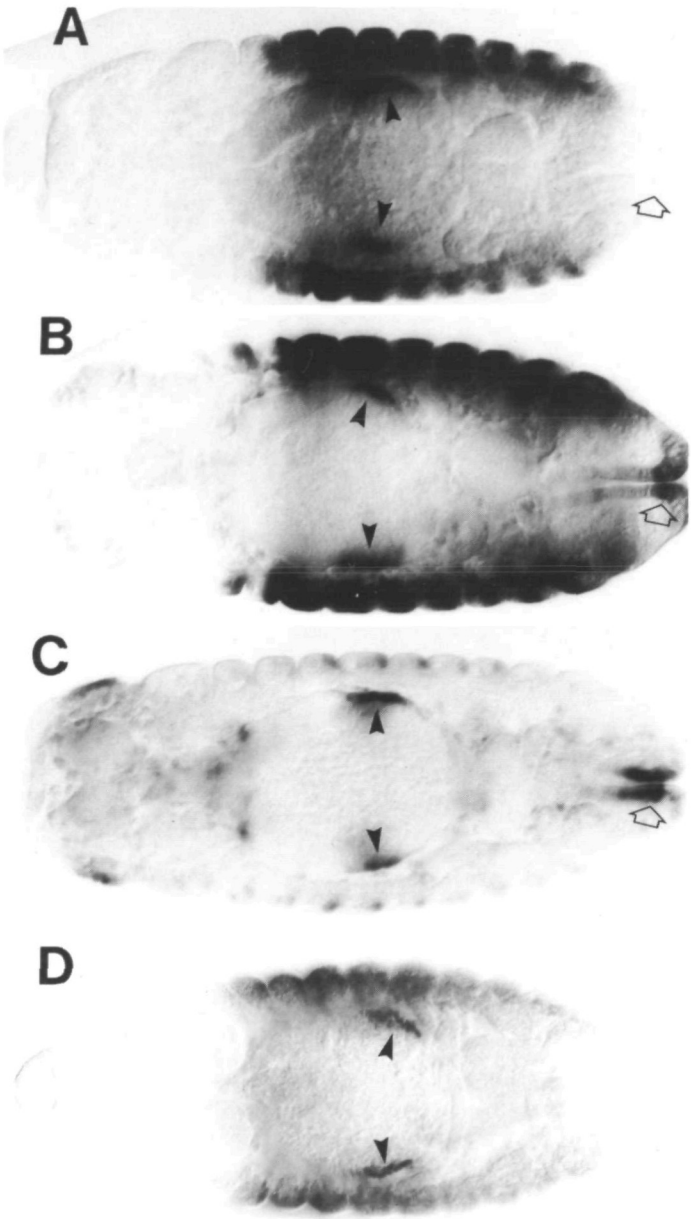


Fig. 7. Expression of UBX and of β-galactosidase in the visceral mesoderm of transformed lines. Horizonal optical sections of stage 14 embryos. The arrowheads mark the PS7 expression in the visceral mesoderm. (A) UBX expression in wild type. (B) β-galactosidase expression in 35UZ. (C) β-galactosidase expression in 5UZ. (D) UBX expression in bxd^{100}. 22UZ (not shown) was expressed similarly in the visceral mesoderm. Immunofluorescence double labels have confirmed that UBX and β-galactosidase expression in transformed lines overlapped (not shown). Note the expression in parts of the hindgut and anal pads in β-galactosidase constructs not seen for UBX (open arrow).

part of the disc (Fig. 9A, B), but was not expressed in other imaginal discs except for some line-specific chromosomal position effects. This expression pattern is like that of UBX in DCR mutants, which eliminate expression in the PS5 compartments of discs but do not affect PS6 disc compartments (Cabrera *et al.* 1985; White and Wilcox, 1985*b*; Botas *et al.* 1988; Little *et al.* 1990).

In 22UZ, β-galactosidase was expressed in complex patterns and at similar levels in all discs and the level of this expression was less than that for 35UZ in PS6. There were common features of expression among different lines, but none of the patterns had any obvious relationship to UBX expression (Fig. 9C). This demonstrates that sequences between 22 and 35 kb upstream of the transcription start site are required for PS6 expression in imaginal cells. The lack of PS6 disc expression in 22UZ is consistent with the loss of imaginal expression of UBX in T3p in bxd^{111} and pbx^{2}

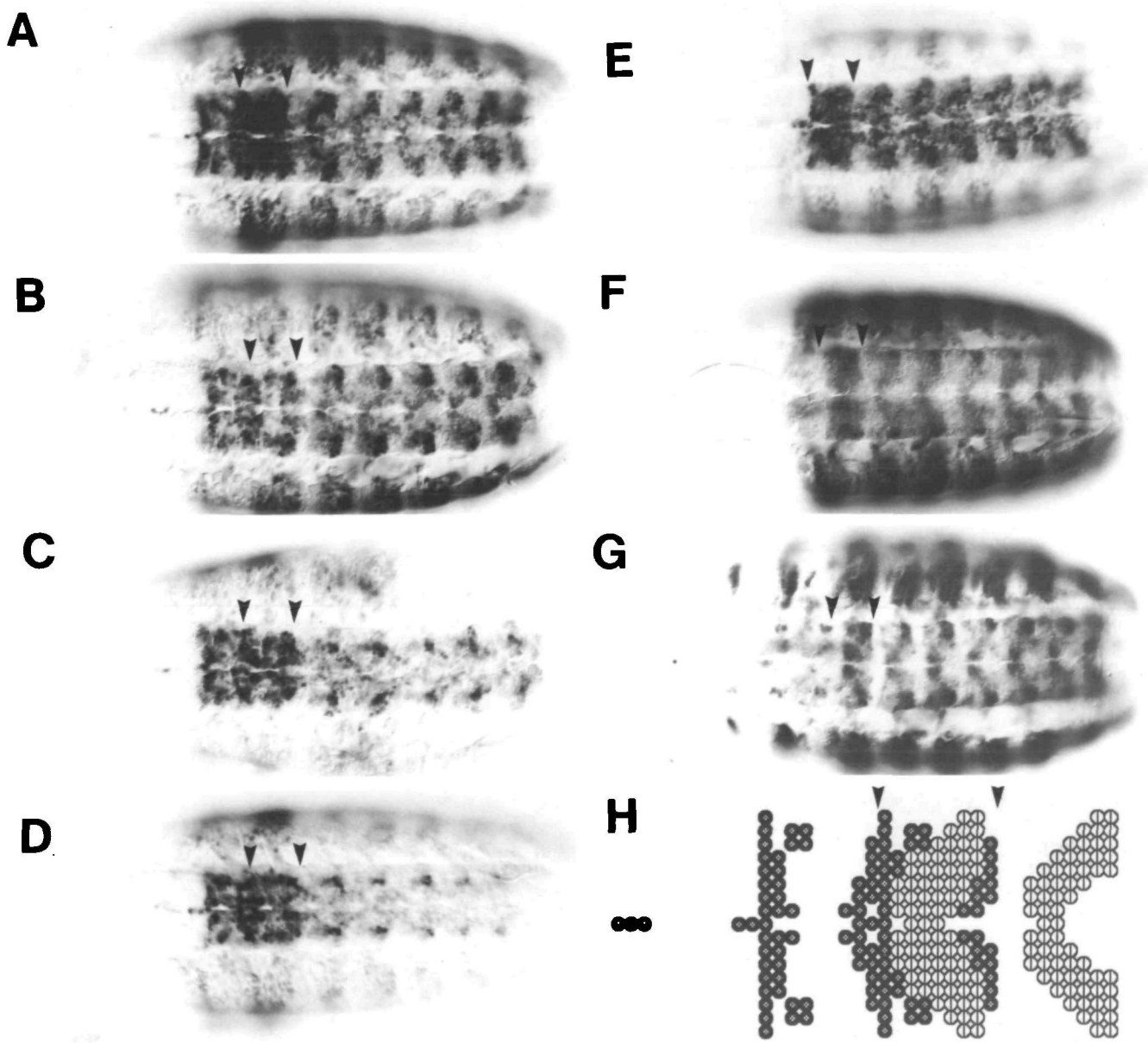


Fig. 8. CNS expression of UBX in wild-type and mutant embryos and of β-galactosidase in 35UZ and 22UZ. Expression was detected by immunohistochemical staining. Horizontal ventral views of stage 14–15 embryos. As these are whole mounts not all parasegments are in exactly the same plane of focus in these photographs. PS6 is delimited by the arrowheads. (A) UBX expression in wild type. (B) UBX expression in bxd^{111} hemizygote. (C) UBX expression in bxd^{106} hemizygote. (D) UBX expression in bxd^{110} hemizygote. (E) UBX expression in abx^{2} hemizygote. (F) β-galactosidase expression in 35UZ. While *Ubx* protein fades away in the epidermis in late embryos, β-galactosidase remains strongly detectable, presumably reflecting its greater stability. The apparent weak expression in 35UZ in PS5 is not in the CNS but is PS6 expression from the overlying epidermis. (G) β-galactosidase expression in 22UZ; note the ectopic expression in PS5 and 4 in both the nervous system and surrounding epidermis. (H) Cartoon of PS4, 5, 6 and 7 showing the division of UBX CNS expression into thoracic (PS5 specific; ●) and abdominal (PS7 specific; ◐) patterns; where these patterns overlap in PS6 the symbol for the abdominal pattern is used. The actual staining pattern observed depends upon the intensity of staining, developmental stage and plane of focus, and is further complicated by cell-to-cell heterogeneity in the levels of staining.

(data not shown), and in pbx^{1} (Cabrera *et al.* 1985; White and Wilcox 1985*b*). These mutations remove DNA located between the 35UZ and 22UZ end points. In contrast, bxd^{68} did not detectably decrease UBX expression in imaginal discs, consistent with the lack of an adult phenotype in T3 in this mutant (Bender *et al.* 1985). 5UZ had weaker imaginal expression which varied from line to line (not shown).

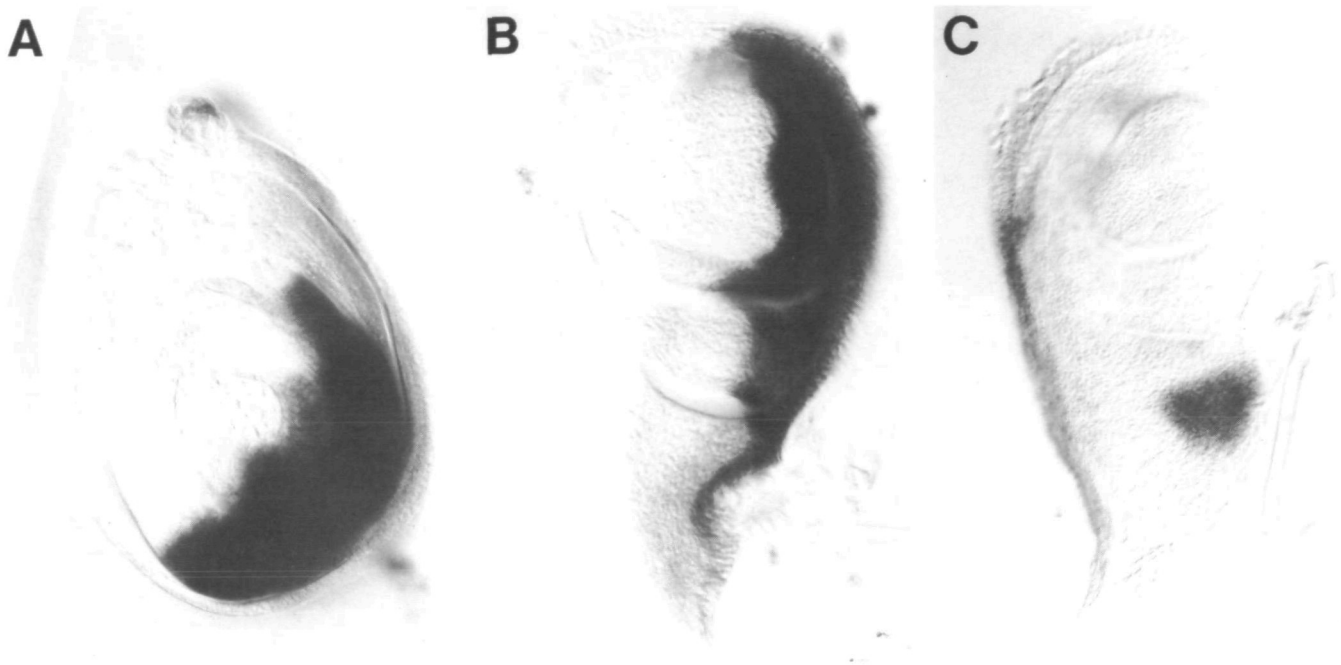


Fig. 9. Imaginal disc expression of β-galactosidase, visualized by activity stains. (*A*) 35UZ T3 leg. Expression was restricted to the posterior compartment. This has been confirmed by double labelling for *en* with fluorescently labelled secondary antibodies (not shown). (*B*) 35UZ haltere disc, expression was restricted to the posterior compartment. The faint expression to the left of the stronger expression was in the peripodial membrane. (*C*) 22UZ haltere disc. This pattern was reproducible from line to line.

Fusion genes have similar patterns of ectopic expression in the head and tail of embryos

All lines expressed β-galactosidase in both the head and tail of the embryo in cells in which UBX expression was not detectable in wild type, *bxd* and *abx* mutants, or in embryos carrying extra chromosomal copies of the BX-C. In stages 9–10, β-galactosidase appeared in the head in a number of cells, many of which, from their position and pattern of staining, appeared to be neural or nerve support cells (Figs 5A and 6). This head expression was detectable in similar cells and at a similar level in all three constructs, although more head cells expressed β-galactosidase in 5UZ (cf. Fig. 6C to A,B). During stage 12, β-galactosidase expression became visible in parts of the hindgut and anal pads in all lines (Fig. 7B,C). Some head and tail expression remained throughout the rest of embryogenesis. This aberrant expression could reflect the absence from the fusion genes of sequences required to repress *Ubx* expression in these positions, or, alternatively, the fortuitous presence of an enhancer of expression for these cells in the vector sequences.

Discussion

The UCR is a large transcriptional regulatory region

Genetic and molecular analysis has indicated that *bxd* mutations alter *Ubx* expression in *cis* (reviewed in Duncan, 1987). It has been argued that this *cis* function must be mediated primarily or exclusively by the UCR DNA, which could act as a giant transcriptional regulatory region (Hogness *et al.* 1985). The ability of 35.4 kb of UCR DNA to confer upon β-galactosidase an expression pattern like that of *Ubx* in DCR mutants throughout development demonstrates that the UCR is a *Ubx* transcriptional regulatory region. The decrease in expression and alterations in pattern observed when only 22.2 kb of UCR DNA are fused to β-galactosidase demonstrate that the UCR is a regulatory region of extraordinary length with important transcriptional regulatory sequences located greater than 22 kb upstream of the transcription start site. The *bxd* phenotypes caused by chromosome breakpoints as far as 35 to 40 kb upstream of the transcription start site suggest that transcriptional regulatory sequences extend at least this far from the promoter. Furthermore, the progressive effects of *bxd* breakpoints distributed throughout the UCR indicate that regulatory elements are spread throughout this control region. This large size can be ascribed, at least in part, to the regulatory complexity of *Ubx* expression. Expression within a cell depends on the developmental stage, tissue, metameric identity and position within a metamere. All this information must be integrated by the *Ubx* control regions to generate the appropriate pattern and level of expression. Similarly, both the P1 (M. Petitt and M. Scott, pers. comm.) and P2 (Boulet and Scott, 1988) promoters of the homeotic gene *Antennapedia* have large transcriptional regulatory regions, and this also appears to be true of the *abdominal A* and *Abdominal B* homeotic genes (Karch *et al.* 1985; Peifer *et al.* 1987)

Early patterning of Ubx

Three classes of zygotic segmentation genes (gap, pair-

rule and segment polarity) execute successive subdivisions of the embryo in the process of metamerization (Nüsslein-Volhard and Wieschaus, 1980; reviewed in Akam, 1987; Scott and Carroll, 1987; Ingham, 1988). Gap genes are expressed and required in broad regions of the embryo, pair-rule genes are expressed and required in alternate metameres, and segment polarity genes are expressed and required in every segment. The patterns of homeotic gene function and expression often respect metameric boundaries, suggesting that the processes of segmentation and segment specification are linked. It has been shown that homeotic gene expression is altered in segmentation gene mutants (Ingham and Martinez-Arias, 1986; Ingham *et al.* 1986; White and Lehmann, 1986; Harding and Levine, 1988; Martinez-Arias and White, 1988; Martinez-Arias *et al.* 1988; Irish *et al.* 1989; Jack *et al.* 1988), indicating that segmentation genes regulate homeotic gene expression. This work presents new results bearing on the role of segmentation genes in the regulation of *Ubx* expression.

It has been reported that UBX protein expression respects the anterior PS6 boundary (Carroll *et al.* 1988). However, using more sensitive detection methods, we detected UBX protein earlier in development and observed that it is initially expressed posterior to the anterior border of PS6. Examination of UBX expression in mutant embryos has shown that the pair-rule gene *fushi tarazu* (*ftz*) is required for this PS6 expression, and it has been proposed that this expression is specified by *ftz* and a locally expressed gap gene in a combinatorial fashion (Ingham and Martinez-Arias, 1986). However, *ftz* expression at this time is strongest at the anterior of the parasegment (p compartment) and fades away posterior to this (Lawrence and Johnston, 1989*a*). Although it is possible that UBX is initially expressed below our sensitivity of detection in the anterior of PS6, this would not alter the fact that there is a lack of correspondance between initial UBX expression (strongest in the posterior of PS6) and *ftz* expression (strongest in the anterior of PS6). This lack of correspondence places constraints on the mechanism of activation of *Ubx* by *ftz*, as the position of UBX expression cannot be explained simply by the position of *ftz* expression. While it is possible that UBX expression is particularly sensitive to combinations of *ftz* and gap gene expression levels, models of this sort are difficult to evaluate because the expression patterns of these genes are dynamic and the lag between their expression and UBX expression is uncertain. Alternatively, other factors may contribute to the determination of the spatial limits of initial *Ubx* activation in PS6.

UBX protein expression occurs in a reproducible spatial and temporal pattern. The pattern of initiation of UBX expression can be divided into five phases, consisting of successive expression in PS6, 8+10+12, 9+11, 5+7 and 13. β-galactosidase expression in 35UZ was similar except that two phases could be subdivided, yielding seven distinguishable phases of expression: PS6, 10+12, 8, 9+11, 5, 7 and finally 13. While the generally high correspondence between UBX and β-galactosidase expression suggests that the difference in the number of distinct phases reflects a difference in the sensitivities of detection, it is also possible that the difference in phases reflects a difference in regulation. In any event, we propose that distinct phases reflect distinct mechanisms of activation; that is, for each phase a unique combination of factors activates *Ubx* expression. Possibly, as suggested for PS6 expression, each phase requires a combination of gap and pair-rule gene products. The identical timing and patterning of expression in PS10 and 12, and then in PS9 and 11, is intriguing, as the two-segment periodicity within a local region would be consistent with this expression being activated by gap and pair-rule genes in combination. While *Ubx* expression is altered in many gap and pair-rule mutants (Ingham *et al.* 1986; Ingham and Martinez-Arias, 1986; White and Lehman, 1986; Martinez-Arias and White, 1988; Irish *et al.* 1989), it has often been difficult to assign the alterations to particular parasegments because of the regulatory interactions among segmentation genes and the gross disruptions in expression pattern observed in many of these mutants. However, careful analysis of the timing and positioning of early expression patterns, including those of *Ubx-lacZ* fusion genes, in mutant embryos may allow further identification of regulatory interactions. Additionally, regulation of *Ubx* by segmentation genes can be fruitfully studied in patches of mutant tissue made by nuclear transplantation (Lawrence and Johnston, 1989*b*).

The persistent pair-rule pattern in 22UZ and strong *bxd* mutants can be attributed to the deletion of sequences with which transcription factors specifying early expression in odd parasegments interact. These sequences are localized to between 22.2 and 29.6 kb upstream of the transciption start site by the early detection of expression in odd parasegments in bxd^{11f}. Detection of this expression in the pbx^2 deletion suggests that another element that can provide this function is located even further upstream. The uncoupling of expression in even and odd parasegments provides evidence that the temporal pattern of *Ubx* expression reflects different activation mechanisms and is consistent with a role for pair-rule genes in the initiation of *Ubx* expression. The later transformation of the pair-rule pattern in 22UZ and *bxd* mutants by the addition of odd parasegmental expression indicates that later expression can be activated by a different mechanism than initial expression. Analogously, spatial patterning of *en* is controlled by two distinct temporal programs, with early expression patterns under control of pair-rule genes that interact with separable sequences in even and odd parasegments, while the later pattern is regulated by segment polarity genes (Dinardo *et al.* 1988).

Organization of Ubx *control regions*

Regulatory elements are spread throughout the UCR, as each of the *bxd* mutants and *Ubx-lacZ* fusions examined have different levels and patterns of ex-

pression. While the subdivision of the UCR is still crude, it appears that there are some tissue-specific regulatory elements. The 5UZ expression pattern indicates that sequences sufficient for correct visceral mesoderm expression, but not for expression in other tissues, are within 5 kb of the transcription start site. 22UZ is expressed in the embryonic ectoderm in a pattern with some similarity to UBX expression. The 17 kb DNA region between the constructs in 5UZ and 22UZ thus contains sequences that allow regulation by some of the factors that govern *Ubx* expression in the ectoderm. Only 35UZ has an expression pattern related to *Ubx* expression in imaginal discs; the 13 kb region between the constructs in 35UZ and 22UZ must interact with factors required for disc expression. The overlap of the pbx^1 and pbx^2 deletions, which both lack T3p disc expression, narrows the location of the imaginal disc element to between 28 and 35 kb upstream of the transcription start site. However, confounding any simple dissection of this element is the observation that fusion of the 13 kb of UCR DNA between the constructs in 22UZ and 35UZ to a heterologous promoter does not generate T3p disc expression (S. Jha and D.S.H., unpubl.). This indicates that T3p disc expression requires interaction between sequences in the −28 to −35 region and regions closer to the promoter, or more specifically, among factors that interact with these regions.

UCR – mechanism of action

What can be the mechanism by which DNA sequences located as far as 35 kb from a promoter are able to alter its transcription? In recent years considerable evidence has accumulated in a variety of systems in favor of the DNA looping model for regulation at a distance. In this model, factors bound to distant sequences interact directly with promoter proximal factors and the intervening DNA is looped out (Ptashne *et al.* 1986; H.-P. Müller *et al.* 1989, and references therein). Transvection between *Ubx* mutations and *Ubx* subfunction mutations (Lewis, 1954) provides genetic evidence favoring a model for the interaction of *Ubx* regulatory sequences by looping. Transvection (reviewed in Judd, 1988; Wu and Goldberg, 1989) is the chromosome-pairing-dependent complementation of mutations in the regulatory sequences of a gene and, at present, can most simply be explained by DNA interactions of the sort envisaged by the looping model. This model, unlike others (reviewed in Ptashne *et al.* 1986), does not require direct DNA linkage between a regulatory sequence and promoter, and thus would allow regulatory sequences to interact with the promoter on the homologous chromosome. Given the complexity both of *Ubx* regulatory sequences and of *Ubx* expression patterns, the model of a loop must be an oversimplification. Elucidation of the actual structure and interactions of *Ubx* control regions will provide a challenge for many years to come, and will likely enhance our knowledge not only of the regulation of *Ubx* expression but of general features of large regulatory regions.

We thank Tom Kornberg, Rob White, and Michael Wilcox for antibodies, Ed Lewis, Welcome Bender, Gerry Rubin and the Indiana Stock Center for *Drosophila* stocks, Vince Pirrotta and Mary Mullins for plasmids, and Yash Hiromi for help with β-galactosidase stains. We are indebted to many past and current members of the group for their advice and assistance throughout this work, and are grateful to Kerry Kornfeld, Juan Botas, Liz Gavis, Cordelia Rauskolb, Sanjaya Jha and Matthew Scott for critical comments on the manuscript. K.D.I. would particularly like to thank Bill Segraves and Howard Lipshitz for advice and encouragement in the early phases of this work. This work was supported by fellowships from the National Science Foundation (K.D.I.), the National Institutes of Health Training Program (K.D.I.), a N.I.N.C.D.S. Physician Scientist Award (S.L.H.) and a National Institutes of Health grant to D.S.H.

References

Akam, M. (1987). The molecular basis for metameric pattern in the *Drosophila* embryo. *Development* **101**, 1–22.

Akam, M. E. and Martinez-Arias, A. (1985). The distribution of *Ultrabithorax* transcripts in *Drosophila* embryos. *EMBO J.* **4**, 1689–1700.

Beachy, P. A., Helfand, S. L. and Hogness, D. S. (1985). Segmental distribution of bithorax complex proteins during *Drosophila* development. *Nature* **313**, 545–551.

Beachy, P. A., Krasnow, M. A., Gavis, E. R. and Hogness, D. S. (1988). An *Ultrabithorax* protein binds sequences near its own and the *Antennapedia* P1 promoters. *Cell* **55**, 1069–1081.

Bellen, H. J., O'Kane, C. J., Wilson, C., Grossniklaus, U., Pearson, R. K. and Gehring, W. J. (1989). P-element-mediated enhancer detection: a versatile method to study development in *Drosophila*. *Genes and Dev.* **3**, 1288–1300.

Bender, W., Akam, M., Karch, F., Beachy, P. A., Peifer, M., Spierer, P., Lewis, E. B. and Hogness, D. S. (1983). Molecular genetics of the bithorax complex in *Drosophila melanogaster*. *Science* **221**, 23–29.

Bender, W., Weiffenbach, B., Karch, F. and Peifer, M. (1985). Domains of *cis*-interaction in the bithorax complex. *Cold Spring Harbor Symp. quant. Biol.* **50**, 173–180.

Bienz, M., Saari, G., Tremml, G., Müller, , Züst, J. and Lawrence, P. A. (1988). Differential regulation of *Ultrabithorax* in two germ layers of *Drosophila*. *Cell* **53**, 567–576.

Bienz, M. and Tremml, G. (1988). Domain of *Ultrabithorax* expression in *Drosophila* visceral mesoderm from autoregulation and exclusion. *Nature* **333**, 576–578.

Botas, J., Cabrera, C. V. and García-Bellido, A. (1988). The reinforcement-extinction process of selector gene activity: a positive feed-back loop and cell–cell interactions in *Ultrabithorax* patterning. *Roux's Arch. Devl Biol.* **197**, 424–434.

Boulet, A. M. and Scott, M. P. (1988). Control elements of the P2 promoter of the *Antennapedia* gene. *Genes and Dev.* **2**, 1600–1614.

Bradford, M. M. (1976). A rapid and sensitive method for the quantitation of microgram quantities of protein utilizing the principle of protein-dye binding. *Anal. Biochem.* **72**, 248–254.

Brower, D. L. (1987). *Ultrabithorax* gene expression in *Drosophila* imaginal discs and larval nervous system. *Development* **101**, 83–92.

Cabrera, C. V., Botas, J. and García-Bellido, A. (1985). Distribution of *Ultrabithorax* proteins in mutants of *Drosophila* bithorax complex and its transregulatory genes. *Nature* **318**, 569–571.

Campos-Ortega, J. A. and Hartenstein, V. (1985). *The Embryonic Development of Drosophila melanogaster*. Berlin: Springer-Verlag.

Canal, I. and Ferrús, A. (1987). The expression of *Ultrabithorax* (*Ubx*) during development of the nervous system of *Drosophila*. *J Neurogenet.* **4**, 161–177.

Carroll, S. B., DiNardo, S., O'Farrell, P. H., White, R. A.

H. AND SCOTT, M. P. (1988). Temporal and spatial relationships between segmentation and homeotic gene expression in *Drosophila* embryos: distributions of the *fushi tarazu*, *engrailed*, *Sex combs reduced*, *Antennapedia*, and *Ultrabithorax* proteins. *Genes and Dev.* **2**, 350–360.

CASANOVA, J., SÁNCHEZ-HERRERO, E. AND MORATA, G. (1985). Prothoracic transformation and functional structure of the *Ultrabithorax* gene of *Drosophila*. *Cell* **42**, 663–669.

DENTE, L., CESARENI, G. AND CORTESE, R. (1983). pEMBL: a new family of single stranded plasmids. *Nucl. Acids Res.* **11**, 1645–1653.

DINARDO, S., KUNER, J. M., THEIS, J. AND O'FARREL, P. H. (1985). Development of embryonic pattern in *D. melanogaster* as revealed by accumulation of the nuclear *engrailed* protein. *Cell* **43**, 59–69.

DINARDO, S., SHER, E., HEEMSKERK-JONGENS, J., KASSIS, J. A. AND O'FARRELL, P. H. (1988). Two-tiered regulation of spatially patterned *engrailed* gene expression during *Drosophila* embryogenesis. *Nature* **332**, 604–609.

DUNCAN, I. (1987). The bithorax complex. *A. Rev. Genet.* **21**, 285–319.

GARCÍA-BELLIDO, A., LAWRENCE, P. A. AND MORATA, G. (1979). Compartments in animal development. *Sci. Am.* **241**, 102–110.

GARCÍA-BELLIDO, A., RIPOLL, P. AND MORATA, G. (1973). Developmental compartmentalisation of the wing disc of *Drosophila*. *Nature New Biol.* **245**, 251–253.

HARDING, K. AND LEVINE, M. (1988). Gap genes define the limits of *Antennapedia* and *Bithorax* gene expression during early development in *Drosophila*. *EMBO J.* **7**, 205–214.

HAZELRIGG, T., LEVIS, R. AND RUBIN, G. M. (1984). Transformation of *white* locus DNA in *Drosophila*: dosage compensation, *zeste* interaction, and position effects. *Cell* **36**, 469–481.

HIROMI, Y., KUROIWA, A. AND GEHRING, W. J. (1985). Control elements of the *Drosophila* segmentation gene *fushi tarazu*. *Cell* **43**, 603–613.

HOGNESS, D. S., LIPSHITZ, H. D., BEACHY, P. A., PEATTIE, D. A., SAINT, R. B., GOLDSCHMIDT-CLERMONT, M., HARTE, P. J., GAVIS, E. R. AND HELFAND, S. L. (1985). Regulation and products of the *Ubx* domain of the bithorax complex. *Cold Spring Harbor Symp. quant. Biol.* **50**, 181–194.

HOOPER, J. E. (1986). Homeotic gene function in the muscles of *Drosophila* larvae. *EMBO J.* **5**, 2321–2329.

INGHAM, P. (1988). The molecular genetics of embryonic pattern formation in *Drosophila*. *Nature* **335**, 25–34.

INGHAM, P., MARTINEZ-ARIAS, A., LAWRENCE, P. A. AND HOWARD, K. (1985). Expression of *engrailed* in the parasegment of *Drosophila*. *Nature* **317**, 634–636.

INGHAM, P. W., ISH-HOROWICZ, D. AND HOWARD, K. R. (1986). Correlative changes in homeotic and segmentation gene expression in *Krüppel* mutant embryos of *Drosophila*. *EMBO J.* **5**, 1659–1665.

INGHAM, P. W. AND MARTINEZ-ARIAS, A. (1986). The correct activation of *Antennapedia* and bithorax complex genes requires the *fushi tarazu* gene. *Nature* **324**, 592–597.

IRISH, V. F., MARTINEZ-ARIAS, A. AND AKAM, M. (1989). Spatial regulation of the *Antennapedia* and *Ultrabithorax* homeotic genes during *Drosophila* early development. *EMBO J.* **8**, 1527–1537.

JACK, T., REGULSKI, M. AND MCGINNIS, W. (1988). Pair-rule segmentation genes regulate the expression of the homeotic selector gene, *Deformed*. *Genes and Dev.* **2**, 635–651.

JOHNSON, F. B. AND KRASNOW, M. A. (1990). Stimulation of transcription by an *Ultrabithorax* protein *in vitro*. *Genes and Dev.* **4**, 1044–1052.

JUDD, B. H. (1988). Transvection: allelic cross talk. *Cell* **53**, 841–843.

KARCH, F., WEIFFENBACH, B., PEIFER, M., BENDER, W., DUNCAN, I., CELNIKER, S., CROSBY, M. AND LEWIŚ, E. B. (1985). The abdominal region of the bithorax complex. *Cell* **43**, 81–96.

KAUFMAN, T., SEEGER, M. A. AND OLSEN, G. (1990). Molecular and genetic organization of the Antennapedia gene complex of *Drosophila melanogaster*. In *Advances in Genetics*, vol. 27 (ed. T. R. F. Wright), pp. 309–362. San Diego: Academic Press.

KORNBERG, T., SIDÉN, I., O'FARRELL, P. AND SIMON, M. (1985). The *engrailed* locus of *Drosophila*: *In situ* localization of transcripts reveals compartment-specific expression. *Cell* **40**, 45–53.

KORNFELD, K., SAINT, R. B., BEACHY, P. A., HARTE, P. J., PEATTIE, D. A. AND HOGNESS, D. S. (1989). Structure and expression of a family of *Ultrabithorax* mRNAs generated by alternative splicing and polyadenylation in *Drosophila*. *Genes and Dev.* **3**, 243–258.

KRASNOW, M. A., SAFFMAN, E. E., KORNFELD, K. AND HOGNESS, D. S. (1989). Transcriptional activation and repression by *Ultrabithorax* proteins in cultured *Drosophila* cells. *Cell* **57**, 1031–1043.

LAWRENCE, P. A. AND JOHNSTON, P. (1989*a*). Pattern formation in the *Drosophila* embryo: allocation of cells to parasegments by *even-skipped* and *fushi tarazu*. *Development* **105**, 761–767.

LAWRENCE, P. A. AND JOHNSTON, P. (1989*b*). Analysis of function of the pair-rule genes *hairy*, *even-skipped* and *fushi tarazu* in mosaic *Drosophila* embryos. *Development* **107**, 847–853.

LAWRENCE, P. A., JOHNSTON, P., MACDONALD, P. AND STRUHL, G. (1987). Borders of parasegments in *Drosophila* embryos are delimited by the *fushi tarazu* and *even-skipped* genes. *Nature* **328**, 440–442.

LEWIS, E. B. (1954). The theory and application of a new method of detecting chromosomal rearrangements in *Drosophila melanogaster*. *Amer. Nat.* **88**, 225–239.

LEWIS, E. B. (1955). Some aspects of position pseudoallelism. *Amer. Nat.* **89**, 73–89.

LEWIS, E. B. (1978). A gene complex controlling segmentation in *Drosophila*. *Nature* **276**, 565–570.

LIPSHITZ, H. D., PEATTIE, D. A AND HOGNESS, D. S. (1987). Novel transcripts from the *Ultrabithorax* domain of the bithorax complex. *Genes and Dev.* **1**, 307–322.

LITTLE, J. W., BYRD, C. A. AND BROWER, D. L. (1990). Effect of *abx*, *bx*, and *pbx* mutations on expression of homeotic genes in *Drosophila* larvae. *Genetics* **124**, 899–908.

MACDONALD, P. M. AND STRUHL, G. (1986). A molecular gradient in early *Drosophila* embryos and its role in specifying the body pattern. *Nature* **324**, 537–545.

MARTINEZ-ARIAS, A., BAKER, N. E. AND INGHAM, P. W. (1988). Role of segment polarity genes in the definition and maintenance of cell states in the *Drosophila* embryo. *Development* **103**, 157–170.

MARTINEZ-ARIAS, A. AND LAWRENCE, P. A. (1985). Parasegments and compartments in the *Drosophila* embryo. *Nature* **313**, 639–642.

MARTINEZ-ARIAS, A. AND WHITE, R. A. H. (1988). *Ultrabithorax* and *engrailed* expression in *Drosophila* embryos mutant for segmentation genes of the pair-rule class. *Development* **102**, 325–338.

MILLER, J. H. (1972). *Experiments in Molecular Genetics*, pp. 352–355. Cold Spring Harbour, New York: Cold Spring Harbor Laboratory.

MITCHISON, T. J. AND SEDAT, J. (1983). Localization of antigenic determinants in whole *Drosophila* embryos. *Devl Biol.* **99**, 261–264.

MÜLLER, H.-P., SOGO, J. M. AND SCHAFFNER, W. (1989). An enhancer stimulates transcription in *trans* when attached to the promoter *via* a protein bridge. *Cell* **58**, 767–777.

MÜLLER, J., THÜRINGER, F., BIGGIN, M., ZÜST, B. AND BIENZ, M. (1989). Coordinate action of a proximal homeoprotein binding site and a distal sequence confers the *Ultrabithorax* expression pattern in the visceral mesoderm. *EMBO J.* **8**, 4143–4151.

MULLINS, M. C., RIO, D. C. AND RUBIN, G. M. (1989). *Cis* acting DNA sequence requirements for P element transposition. *Genes and Dev.* **3**, 729–738.

NÜSSLEIN-VOLHARD AND WIESCHAUS, E. (1980). Mutations affecting segment number and polarity in *Drosophila*. *Nature* **287**, 795–801.

O'CONNOR, M. B., BINARI, R., PERKINS, L. A. AND BENDER, W. (1988). Alternative RNA products from the *Ultrabithorax* domain of the bithorax complex. *EMBO J.* **7**, 435–445.

PATEL, N. H., MARTIN-BLANCO, E., COLEMAN, K. G., POOLE, S. J., ELLIS, M. C., KORNBERG, T. B. AND GOODMAN, C. S. (1989).

Expression of *engrailed* proteins in Arthropods, Annelids, and Chordates. *Cell* **58**, 955–968.

PEIFER, M. AND BENDER, W. (1986). The *anterobithorax* and *bithorax* mutations of the bithorax complex. *EMBO J.* **5**, 2293–2303.

PEIFER, M., KARCH, F. AND BENDER, W. (1987). The bithorax complex: control of segmental identity. *Genes and Dev.* **1**, 891–898.

PIRROTTA, V. (1988). Vectors for P-mediated transformation in *Drosophila*. In *Vectors: A Survey of Molecular Cloning Vectors and their Uses* (eds. R. L. Rodriguez and D. T. Reinhardt), pp. 437–456. Boston: Butterworths.

PTASHNE, M. (1986). Gene regulation by proteins acting nearby and at a distance. *Nature* **322**, 697–701.

ROBERTSON, H. M., PRESTON, C. R., PHILLIS, R. W., JOHNSON-SCHLITZ, D. M., BENZ, W. K. AND ENGELS, W. R. (1988). A stable genomic source of P element transposase in *Drosophila melanogaster*. *Genetics* **118**, 461–470.

SCOTT, M. P. AND CARROLL, S. B. (1987). The segmentation and homeotic gene network in early *Drosophila* development. *Cell* **51**, 689–698.

SCOTT, M. P., TAMKUN, J. W. AND HARTZELL, G. W. III (1989). The structure and function of the homeodomain. *BBA Rev. Cancer* **989**, 25–48.

SIMON, J., PEIFER, M., BENDER, W. AND O'CONNOR, M. (1990). Identification of parasegmental control elements within the bithorax complex of *Drosophila melanogaster*. *EMBO J.* **9**, (in press).

SPRADLING, A. C. AND RUBIN, G. M. (1982). Transposition of cloned P elements into *Drosophila* germ line chromosomes. *Science* **218**, 341–347.

STELLER, H. AND PIRROTTA, V. (1985). A transposable P vector that confers selectable G418 resistance to *Drosophila* larvae. *EMBO J.* **4**, 167–171.

TEUGELS, E. AND GHYSEN, A. (1985). Domains of action of bithorax genes in *Drosophila* central nervous system. *Nature* **314**, 558–561.

THUMMEL, C. S., BOULET, A. M. AND LIPSHITZ, H. D. (1988). Vectors for *Drosophila* P element-mediated transformation and tissue culture transfection. *Gene* **74**, 445–450.

THUMMEL, C. S., BURTIS, K. C. AND HOGNESS, D. S. (1990). Spatial and temporal patterns of *E74* transcription during *Drosophila* development. *Cell* **61**, 101–111.

WHITE, R. A. H. AND LEHMANN, R. (1986). A gap gene, *hunchback*, regulates the spatial expression of *Ultrabithorax*. *Cell* **47**, 311–321.

WHITE, R. A. H. AND WILCOX, M. (1984). Protein products of the bithorax complex of *Drosophila*. *Cell* **39**, 163–171.

WHITE, R. A. H. AND WILCOX, M. (1985*a*). Distribution of *Ultrabithorax* proteins in *Drosophila*. *EMBO J.* **4**, 2035–2043.

WHITE, R. A. H. AND WILCOX, M. (1985*b*). Regulation of the distribution of *Ultrabithorax* proteins in *Drosophila*. *Nature* **318**, 563–567.

WU, C.-T. AND GOLDBERG, M. L. (1989). The *Drosophila* zeste gene and transvection. *Trends Genet.* **5**, 189–194.

ZUKER, C. S., COWMAN, A. F. AND RUBIN, G. M. (1985). Isolation and structure of a rhodopsin gene from *D. melanogaster*. *Cell* **40**, 851–858.

(*Accepted 6 November 1990*)