Homeotic control in *Drosophila*; the *scabrous* gene is an *in vivo* target of *Ultrabithorax* proteins

Yacine Graba, Denise Aragnol,
Patrick Laurenti, Véronique Garzino¹,
Dominique Charmot, Hélène Berenger and
Jacques Pradel

Laboratoire de Génétique et de Biologie Cellulaires, C.N.R.S., Centre
Universitaire de Marseille-Luminy, Case 907, 13288 Marseille cedex
9, France

¹Present address: Institute of Zoology, Laboratory of Neurobiology,
Rheinsprung 9, Basel 4051, Switzerland

Communicated by A.Ghysen

The regulatory functions of transcription factors encoded by the *Ultrabithorax* (*Ubx*) gene initiate genetic programmes essential for segmental identity and morphogenesis in *Drosophila*. Based on the formation of DNA–protein adducts in intact nuclei and immunoselection procedure, we cloned genomic targets for *Ubx* proteins. One clone was studied in detail. It encompasses parts of the last intron and exon of the *scabrous* (*sca*) gene, which encodes a secreted protein involved in cellular communication during neurogenesis. Five motifs, presenting the ATTA core, which is shared by most homeodomain binding sites, were found in the nucleotide sequence of this clone. We detail here the dynamic pattern of *sca* transcript accumulation during embryogenesis and show that mutation of *Ubx* results in the ectopic transcription of *sca* in the first abdominal segment. We propose that a direct interaction of *Ubx* with *cis-acting* elements in *sca* negatively regulates the gene. Transcript localization in several combinations of deficiencies in the Bithorax complex (BX-C) indicates that *sca* is downregulated by *abdominal A* (*abdA*) and *Abdominal B* (*AbdB*), and suggests that it is a common target of the three genes of BX-C.

Key words: Bithorax complex/*Drosophila*/neurogenesis/*scabrous* transcription pattern/*Ultrabithorax* target genes

Introduction

Genetic control of metamerization in *Drosophila* is now well understood (Akam, 1987; Nüsslein-Volhard *et al.*, 1987; Ingham, 1988). A hierarchy of genetic modules, comprising maternal, segmentation and homeotic genes, are sequentially deployed to specify the elements of metameric pattern at progressively finer and more detailed levels. The temporal and spatial ordering of gene expression within this hierarchy results from regulatory interactions: the products of maternally expressed genes regulate segmentation gene activity, the products of which then control homeotic gene expression. In addition, interactions also occur between genes of a same class within the hierarchy to define exclusive domains of expression (for reviews see Ingham and

Martinez-Arias, 1992; St Johnston and Nüsslein-Volhard, 1992; McGinnis and Krumlauf, 1992). Therefore after 5–6 h of development, each metameric unit specifies its own identity by expressing a unique combination of homeotic genes. Since homeoproteins are thought to function as transcription factors (Gehring, 1987; Scott *et al.*, 1989), each metamer will enter a specific morphogenetic pathway depending on the target genes whose activity is regulated by its set of homeoproteins. How this specificity is achieved remains obscure. It presumably results from a complex interplay of differential affinity and competition of homeoproteins for common targets (Morata and Struhl, 1990), subtle DNA binding preferences (Dessain *et al.*, 1992) and interactions with other proteic cofactors (Peifer and Wieschaus, 1990; Fitzpatrick *et al.*, 1992; Smith and Johnson, 1992). Only few candidate effector genes have so far been identified. Understanding the specificity of action of homeotic genes in the whole animal obviously requires the identification of additional 'downstream' genes. This is now an open and actively investigated field (Gould *et al.*, 1990; Wagner-Bernholz *et al.*, 1991; Andrew and Scott, 1992).

Ultrabithorax (*Ubx*) homeotic gene specifies and maintains the identity of parasegments (PS) 5 and 6 and is also required for the correct formation of the midgut. *Ubx* produces a series of protein isoforms due to alternative splicing and post-translational maturation (Busturia *et al.*, 1990; Gavis and Hogness, 1991), which act, probably in concert with positive or negative co-factors, either as transcriptional activators or repressors (Johnson and Krasnow, 1990). It appears highly probable that *Ubx* proteins are involved in the regulation of a number of downstream genes. Genetic methods identify few putative targets, for example *decapentaplegic* or *wingless*; both of which are mammalian growth factor homologues expressed in the visceral mesoderm (Immerglück *et al.*, 1990; Reuter *et al.*, 1990). Gould *et al.* (1990) have developed a molecular approach based on the direct isolation of genomic sequences associated with *Ubx* proteins in native embryonic chromatin. The attempt has been successful and two transcription units have been identified as putative *Ubx* targets. We have developed a parallel approach based on the stabilization, by UV light irradiation, of protein–DNA interactions existing in intact embryonic nuclei in order to immunopurify and clone genomic targets for *Ubx* proteins. We characterize one clone in detail. It is located within *scabrous* (*sca*), a gene that produces a secreted protein involved in cellular communication during neurogenesis. A *sca* function in lateral inhibition within the developing nervous system has been inferred from detailed analyses of its expression pattern in eye and wing imaginal discs and phenotypes of deficient animals (Baker *et al.*, 1990; Mlodzik *et al.*, 1990). Here, we analyse the dynamic changes of the *sca* transcription pattern during early embryogenesis and show that expression

in abdominal segments is downregulated by *Ubx* and also by the two other genes of the Bithorax complex, *abdominalA* (*abdA*) and *AbdominalB* (*AbdB*).

Results

Immunopurification and cloning of *Ubx* proteins genomic targets

The procedure that was used for purifying genomic targets of *Ubx* proteins is outlined in Figure 1A. Our major concern has been to preserve the DNA–protein interactions existing

Fig. 1. Cloning the genomic target of *Ubx* proteins. **(A)** The stepwise procedure. **(B)** UV irradiation of isolated nuclei links the *Ubx* protein to DNA. Increasing amounts (1, 4 and 20 μ g) of CsCl, purified DNA from irradiated (lanes 2–5) or non-irradiated nuclei (lane 1), were dot blotted onto nitrocellulose. Probing with monoclonal antibody FP3.38 (lanes 1 and 2) clearly shows that UV irradiation links the *Ubx* protein to the DNA. Lanes 3–5 are positive and negative controls; lane 3 shows that proteinase K abolishes the reaction with FP3.38; lane 5 shows that probing with other antibodies indicates that RNA polymerase 2B is cross-linked to DNA, but not an antigen against an extracellular matrix component (lane 4), monoclonal antibody RD3 (Garzino *et al.*, 1989), also binds DNA upon UV irradiation. **(C)** Immunoprecipitation of DNA–*Ubx* protein adducts. DNA (2 mg) purified from UV irradiated nuclei was mechanically sheared by sonication and the DNA–*Ubx* protein adducts immunoprecipitated. Sub-fractions of the immunoprecipitate (1/100, 1/50, 1/10, 1/4 and 1/2) were dot blotted onto nitrocellulose and probed with FP3.38.

in the developing embryo prior to selecting the relevant DNA fragments with a specific antibody against *Ubx* proteins. A first important point has been to isolate embryonic nuclei in conditions that prevent the release of nuclear proteins. This release occurs when conventional methods are used (Paine *et al.*, 1983; Garzino *et al.*, 1987), which could, in turn, result in changing the binding pattern of chromatin associated proteins. Stabilization of DNA–protein interactions was performed by irradiating nuclei with UV light, following a protocol described by Gilmour and Lis (1985). UV cross-linking is a powerful method that has been successfully used to characterize the *in vivo* DNA-binding activity of several proteins (Gilmour *et al.*, 1986; Champlin *et al.*, 1991). Under a controlled irradiation, only those proteins that are in direct contact with DNA can form covalent adducts (Chatterjee *et al.*, 1988) and remain associated to DNA during CsCl gradient purification, the next step in our protocol. The DNA fraction obtained from the CsCl gradient was then digested with endonuclease *Eco*RI. The DNA fragments covalently linked to *Ubx* proteins were immunoprecipitated by FP3.38, a monoclonal antibody that recognizes all the *Ubx* protein isoforms (White and Wilcox, 1984). Using an end-labelled aliquot fraction of the *Eco*RI digest as a tracer, we estimated that about 5×10^{-5} of the total genomic DNA was recovered as *Ubx* protein associated DNA. After digestion of covalently attached proteins with proteinase K, the resulting DNA fragments were cloned into λ gt10. The cloning efficiency was rather poor since only 2×10^4 recombinant phages were obtained from 50 ng of immunopurified fragments.

Figures 1B and C represent an analysis of the efficiency of UV cross-linking and immunopurification steps. The DNA sample purified from nuclei, which had previously been UV irradiated (Figure 1B, lane 2) or not (lane 1), was dot-blotted onto nitrocellulose and probed with the antibody FP3.38. Clearly, UV light irradiation promotes the formation of *Ubx* protein–DNA adducts. Proteinase K treatment abolished the reaction (lane 3); an antibody against an extracellular antigen did not react (lane 4); RNA polymerase 2B was also cross-linked to DNA and was revealed by a specific antibody (lane 5). Figure 1C positively shows that *Ubx* protein–DNA adducts were immunoprecipitated after the CsCl gradient step.

We assayed a series of 17 clones randomly selected from the library for their ability to bind *Ubx* protein *in vitro*. DNA fragments were incubated with *Ubx* protein made in *Escherichia coli* and immunoprecipitated with antibody FP3.38. Binding efficiency was estimated, in the presence of high concentrations of non-specific competitive DNA, by the value of the ionic strength that releases the DNA fragment from *Ubx* protein. Figure 2 shows the result of a binding test performed on a mixture of four clones. Two clones, Cu4 and Cu9, did not bind *Ubx* protein in our assay conditions. Selective *Ubx* protein binding was observed with the two other clones: Cu6 exhibited a strong affinity as the interaction still occurred at 1 M NaCl; Cu8, released at 0.4 M NaCl, showed a more reduced affinity. There is no correlation between the length of the insert and the *Ubx in vitro* binding activity. Over the 17 randomly selected clones, 7 clones, sizing from 150 bp to 1.5 kb, were able to bind *Ubx* protein *in vitro*. Five of the seven clones reacted on Northern blots. The distributions of their cognate mRNAs were assayed by hybridization *in situ* on whole-mount embryos. For three clones, Cu5, Cu8 and Cu83, the results are clearly consistent

with a regulation by *Ubx* (data not shown). We will focus in this paper on clone Cu8.

Based on these controls, one can assume that while containing a number of irrelevant sequences, the library constructed from genomic fragments associated with *Ubx*

gene products in a context close to the *in vivo* situation was significantly enriched in fragments containing binding sites for *Ubx* proteins. It appears obvious, however, that all of the targets have not been cloned, since using specific probes for the promoter P1 of *Antenapedia* (*Antp*) (Beachy *et al.*, 1988) and for clones 35 and 48 (Gould *et al.*, 1990), we failed to detect these putative targets in the library.

Clone Cu8 is an *EcoRI* fragment of the scabrous gene

The clone Cu8 maps to 49 D/E cytological region and hybridizes to a 3 kb mRNA on Northern blots (data not shown). These data fit the cytological location and transcript size of the *sca* gene (Mlodzik *et al.*, 1990). Comparison of the Cu8 nucleotide sequence with available sequence data of *sca* (the coding region and sequences in the untranslated 5' part, but not introns, are published or available in databanks; Baker *et al.*, 1990; Mlodzik *et al.*, 1990) established that the clone Cu8 (480 bp) is an *EcoRI* genomic fragment situated in the 3' part of the gene. This fragment contains 123 nucleotides of the third intron and 357 nucleotides of the fourth exon (Figure 3). The exonic sequence almost matches 357 nucleotides in Cu8. There is one mismatch, shown in Figure 3, which changes one histidine for one aspartic acid in Cu8, which presumably reflects a strain polymorphism. The analysis of the Cu8 sequence revealed the presence of five motifs (underlined in Figure 3), presenting the ATTA core which is found in most homeodomain binding sites (Treisman *et al.*, 1992). Two of the three motifs lying in the intronic part, TCAATTATTT and TTTATTAAAT, resemble the consensus TCAATTAAAT for homeoproteins binding (Desplan and al, 1988; Hoey and Levine, 1988; Thali *et al.*, 1988) and the motif TCAATTAAAT, identified by Gould

Fig. 2. Selective binding to *Ubx* protein *in vitro*. *Ubx* protein made in *E. coli* was incubated with labelled DNA fragments from four clones of the library (Cu4, Cu6, Cu8 and Cu9) in the presence of 100- and 2.5×10^3 -fold excess of herring sperm DNA and poly dI-dC respectively, and immunoprecipitated by FP3.38 coupled to trisacryl beads. The DNA fragments were then eluted by increasing salt concentration. Lane 1, input DNA; lanes 2, 3 and 4, 0.2 M NaCl washes; lane 5, 0.3 M NaCl; lane 6, 0.4 M NaCl; lane 7, 0.5 M NaCl; lane 8, 1 M NaCl. Clones Cu6 and Cu8 showed selective binding. Sizes are given in kb on the left.

Fig. 3. Clone Cu8 is an *EcoRI* fragment of the *sca* gene. Nucleotide sequence revealed a quasi-perfect match of the last 357 nucleotides in Cu8 (capital letters) with the first 357 nucleotides of the fourth exon of *sca* (Baker *et al.*, 1990). Nucleotide mismatches, shown by an asterisk, change G-C in *sca* for C-G in Cu8, resulting in the replacement of a histidine by an aspartic acid. Motifs presenting the ATTA core are underlined. In the 123 nucleotides intronic part of Cu8 (small letters), two motifs resembling the TCAATTAAAT consensus are also underlined. *EcoRI* restriction map and *sca* gene structure are from Mlodzik *et al.* (1990).

et al. (1990) in clone 48, a putative *in vivo* target of *Ubx* proteins.

Seven additional ATTA core motifs lie in *sca*. Four of these motifs sit in the 5' leader region near the TATA box, the remaining three being scattered within the coding part of the gene.

sca expression in wild type embryo

Mlodzik *et al.* (1990) provided a much detailed analysis of *sca* transcription during imaginal disc development. The embryonic pattern was also briefly considered as revealed by *in situ* hybridization of embryonic sections and anti- β -galactosidase immunohistochemistry in a line carrying a *P(lacZ)* insertion close to the 5' end of *sca*. Nothing in this description of the embryonic pattern was suggestive for a possible homeotic control along the anteroposterior axis. Since our molecular data suggested a transcriptional control by *Ubx* involving sequences lying in the 3' part rather than in the 5' end of the gene, we decided for a re-examination

of *sca* expression using *in situ* hybridization on whole-mount embryos.

sca transcripts are first detected in all cells at cellular blastoderm (Figure 4A; stage 5, according to Campos-Ortega and Hartenstein, 1985). During subsequent embryogenesis the *sca* expression pattern shows very dynamic changes. By the end of stage 5, staining is intensified in some regions of the embryo, defining a large anterior stripe in the head and seven ventrolateral stripes in the presumptive trunk neuroectoderm (Figures 4B and C). By contrast, *sca* transcripts are no longer detected at anterior and posterior poles of the embryo nor in the presumptive mesoderm (ventral views in Figures 4D and E). During early germ band extension, the staining becomes broad and thick and progressively invades the ventrolateral ectoderm (Figures 4F and G). By the end of stage 8, before the completion of germ band extension, the gene is preferentially expressed in 14 segmentally repeated cell clusters in the neuroectoderm (Figure 4H).

Fig. 4. The distribution of *sca* transcript during early embryogenesis. Embryos (anterior to the left, ventral side down in lateral views) were staged according to Campos-Ortega and Hartenstein (1985). *sca* expression was detected in whole-mount embryos with digoxigenin labelled probe. (A) Cellular blastoderm (stage 5), lateral view showing a transient uniform distribution of *sca* transcript. (B) Same orientation for a slightly older embryo (late stage 5) displaying the seven stripes with a pair rule-like and dorsoventral modulated pattern. Note also the large anterior stripe in the procephalon. (C and D) Early gastrula (stage 6), ventral and lateral views. The surface of the embryo is out of focus in (C). *sca* transcript is no longer visible at both anterior and posterior poles. The ventrolateral stripes are unequal in wideness and intensity. Notice in the surface view in (D) that the mesoderm anlage is devoid of labelling. (E) Late gastrula (stage 7), ventral view. The arrowhead points to a higher density of transcript just behind the posterior border of cephalic furrow. (F and G) Rapid germ band elongation (early stage 8), lateral and ventral views, respectively. Transcripts are visible in all the cells of the neurogenic ectoderm with a modulation in alternating domains of high and low accumulation along the anteroposterior axis. (H) Slow germ band elongation (stage 8), lateral view. Note the segment polarity-like pattern in 14 clusters of cells where *sca* transcript preferentially accumulates. Bar represents 50 μ m.

The development of the CNS starts with the separation of neural progenitor cells from the neurogenic region of the ectoderm. From stages 9–11, delamination of neuroblasts proceeds discontinuously in three pulses giving rise to three subpopulations that are arranged in parallel rows along the dorsoventral axis (Hartenstein and Campos-Ortega, 1984). During this process, *sca* is strongly expressed in the three rows of neuroblasts and at a reduced level in cells which differentiate as epidermoblasts (Figures 5A–C). This expression in neuroblasts occurs transiently; some time after delamination, *sca* is no longer transcribed in the segregated neuroblasts that have moved inward from the surface of the embryo (stage 11, Figure 5D).

The progenitors to sensory organs, the sensory mother cells (SMC), represent a subset of cells from the lateral ectoderm that adopt a neural fate while neighbouring cells take on an epidermal fate (Ghysen and Dambly-Chaudière, 1989). In the lateral ectoderm of early stage 10 embryo, *sca* expression is initiated in one cell per PS (Figure 5B). On

the basis of their position, just posterior to the parasegmental grooves that form at that time, these cells could correspond to the first appearing SMC, the P cells (Ghysen and O’Kane, 1990). At the end of germ band elongation (early stage 11), transcripts are visible in an increased number of SMC (Figure 5E). Staining of cell clusters in gnathal segments and procephalon is also consistent with an expression in the developing sensory organs in the head. Just as in the developing CNS, *sca* is transiently expressed during PNS development. By early stage 12, the precursors to PNS in the trunk region are dividing actively (Bodmer *et al.*, 1989) giving rise to groups of related cells, the primordia of sensory organs. *sca* is not expressed (or at a low level) in those cells (Figure 5F).

Progenitors for other neural lineages also express the *sca* gene. Transcripts are detected within the midline progenitor cells (arrow in Figure 5D; Thomas *et al.*, 1984). In addition, during early stage 12, transcripts are visible in four rows of cells that paralleled the midline of the embryo, having

Fig. 5. The distribution of *sca* transcript in the elongating germ band. (A) *sca* RNA is distributed in the neuroblasts that delaminate from the ventral ectoderm and in the procephalon at stage 9. (B) In the early stage 10 embryo, this distribution persists. At that time, accumulation of transcript is also visible in PNS progenitor cells that appear as single cells or doublets in the lateral ectoderm of each trunk segment (arrowhead). (C) The ventral view of a late stage 10 embryo shows expression in the six longitudinal rows of neuroblasts that parallel the midline and in precursors of PNS in the lateral ectoderm. (D and E) In stage 11 embryos, the segregated neuroblasts, which have moved inward from the surface, are devoid of transcript. In contrast, *sca* is expressed in additional precursor cells to PNS in each segment, A8 excepted. Note expression in midline progenitor cells (arrowhead in the ventral view D) and in cell clusters of the procephalon, which presumably correspond to antennomaxillary sense organs and gnathal neurogenic region. (F, G and H) In stage 12 embryos, cells of the developing PNS are no longer stained. The label in the head is restricted dorsolaterally to the primordia of optic lobe and procephalic neurogenic region (arrows in F; also visible in the dorsal view in H). Note the expression of *sca* in four rows of cells along the midline (arrows in G); three cells per hemisegment are labelled. Moreover, a segment to segment modulation occurs at this stage. Arrowheads in (F) point to *sca* RNA accumulation within small clusters of ectodermal cells in each of the three thoracic segments. Bar represents 50 μ m.

a segmental repeated pattern (Figures 5G–H). They could correspond either to commissural neurons pioneer cells (Patel *et al.*, 1987), which will interconnect the two halves of the CNS (Klambt *et al.*, 1991), or to precursors of the ventralmost groups of the PNS (Bodmer *et al.*, 1989).

At the same time, *sca* expression appears in a segment to segment modulated pattern. Small clusters of 8–10 ectodermal cells are stained in the three thoracic but not in abdominal segments (Figures 5F–G). According to their size and shape, position on the ventral side of the ectoderm, segmental specificity and behaviour in mutant background (see below), these cells might correspond to sense organ precursors of the thoracic segments. Labelled cell clusters do not increase in size as the development proceeds; rather, *sca* transcript accumulation in these clusters (as well as in the four rows of cells that parallel the midline), occurs transiently during stage 12. Transcription is no longer detected in the trunk until stage 17, just before the hatching of the embryo, where a low level of *sca* expression is visible in the ventral nerve cord (not shown).

sca expression in *Ubx* and *Antp* mutants

By the end of stage 11 when the germ band is fully extended and by early stage 12 when the germ band begins retraction, i.e. precisely when the expression of homeotic genes is functionally relevant (e.g. Carroll *et al.*, 1988), *sca* transcription pattern shows segment to segment modulation. Transcripts are detected in small clusters of ectodermal cells in thoracic but not in abdominal segments. By this time, *Ubx* proteins have accumulated following a complementary pattern, mainly in the first abdominal segment (A1) and abdominal segments 2–7 (A2–7), weakly in the third thoracic segment (T3) and not in the first and second thoracic segments (T1–2). These data, when taken together with the molecular evidence presented earlier showing that *Ubx* proteins bind the 480 bp *EcoRI sca* fragment, strongly suggest a negative regulation by *Ubx* proteins.

The distribution of *sca* transcript was studied in embryos homozygous for a deficiency of the *Ubx* gene, *Ubx*^{9.22} (Kerridge and Morata, 1982). No modification in the wild type pattern is observed in these embryos before stage 12. As the germ band begins to retract, however, *sca* becomes ectopically expressed in A1 (Figure 6). Temporally, the ectopic expression in A1 is seen as early and ends at the same time as the expression in thoracic segments. An alternative explanation for the negative control by *Ubx* in the abdomen could be positive regulation by *Antp* in the thorax. *Antp* protein indeed accumulates in the three thoracic, but not in abdominal, segments; moreover, in *Ubx* mutant the ectopic expression of *Antp* that occurs in A1 (Hafen *et al.*, 1984) would explain the ectopic expression of *sca*. When analysing the *sca* transcription pattern in *Antp*^{W10}, a mutant deficient for the *Antp* gene (Wakimoto *et al.*, 1984), we failed to detect any change with regard to the wild type pattern (not shown). These results indicate that *Ubx* functions to repress *sca* in A1 of the wild type embryo. The ectopic expression occurs in a small cluster of ectodermal cells that mimics the clusters seen in thoracic segments. This is consistent with the phenotype of *Ubx* mutants, in which A1 takes on a thoracic identity (Lewis, 1978; Hayes *et al.*, 1984).

Fig. 6. Ectopic expression of *sca* in embryo deficient for *Ubx*. Lateral views of embryos at the onset of germ band retraction. The arrowheads indicate the first abdominal segment in the two panels. The wild type expression occurs in small cluster of cells in the three thoracic segments (A) while ectopic expression of *sca* is observed in the first abdominal segment of mutants deficient for *Ubx* (B). Bar represents 50 μ m.

sca expression in other BX-C mutants

Since the repression of *sca* was only raised in the first abdominal segment in a mutant deficient for *Ubx*, we looked for a role of the other homeotic genes of BX-C in regulating the *sca* transcription in abdominal segments A2–9. *abdA* is expressed from A2–8 (Harding *et al.*, 1985) and *AbdB* from A5–9 (Celniker *et al.*, 1989). In a *Ubx abdA* double mutant, Df(3R)109 (Lewis, 1978), *sca* is ectopically expressed not only in A1 but also in A2–7 (Figures 7A, C and E). In a mutant deficient for the entire BX-C, Df(3R)P9 (Lewis, 1978), the ectopic expression occurred in all of the abdominal segments (Figures 7B, D and F). Accordingly, we propose that in wild type embryos the *abdA* and *AbdB* proteins repress *sca* transcription in segments A2–7 and A8–9, respectively.

Furthermore, we observed that in a single mutant for either *abdA* (*abdA*^{MX1}; Sanchez-herrero *et al.*, 1985) or *AbdB* (*AbdB*^{S4}; Casanova *et al.*, 1986), the wild type pattern of *sca* transcription is not altered (not shown). The simplest explanation is that the ectopically expressed *Ubx* and *abdA* gene products (Struhl and White, 1985) are able to silence *sca* in A2–7 of *abdA*^{MX1} and in A8–9 of *AbdB*^{S4} embryos, respectively. This strengthens the previous proposal that the three genes of BX-C could independently repress *sca* in the abdomen.

Discussion

This paper reports the strategy we used to isolate and clone *EcoRI* genomic fragments associated with *Ubx* proteins in the developing embryo. The procedure for the recovery of DNA fragments bound to a protein of interest is quick and easy to perform and needs only the use of specific antibodies. The method may be thus generally applicable for identifying downstream targets of transcription factors. There is, however, two obvious limitations that result from the UV

Fig. 7. Ectopic expression of *sca* in embryos deficient for BX-C genes. Lateral (A and B), ventral (C and D) and dorsal (E and F) views of stage 12 embryos. *sca* is ectopically expressed in segment A1–A7 of embryos deficient for *Ubx* and *abdA* (panels A, C and E) and in segment A1–A9 in mutants deficient for the three genes of BX-C. Ectopic expression occurs in small clusters of ectodermal cells that resemble the clusters seen in thoracic segments. Bar represents 50 μ m.

light irradiation. First, this treatment promotes the linkage of an average of one protein molecule to DNA every 60 kb (Gilmour and Lis, 1985), which means that the probability for a given protein to become covalently attached to its target DNA is weak. Secondly, the cloning efficiency of recovered DNA is manifestly poor, which could be due to the incomplete release of amino acid residues upon proteinase K digestion and/or to DNA damages induced by UV light.

A library of *Ubx* genomic targets

Due to the relative lack of specificity of DNA–homeoprotein interactions, it appears rather hazardous to attempt the cloning and identification of genes acting downstream to homeotic genes using *in vitro* reconstituting assays. Two promising approaches have been recently deployed aiming the enhancer detection of target genes of *Antp* (Wagner-Bernholz *et al.*, 1991) and the immunoselection of *in vivo* targets of *Ubx* (Gould *et al.*, 1990). The main difference between the approach developed by Gould *et al.* (1990) and ours rests on the procedure used to preserve the interactions existing *in vivo*. Gould *et al.* (1990) prepared restriction fragments from native chromatin isolated at low ionic strength. Alternatively, we performed UV cross-linking experiments to freeze the DNA–protein interactions existing in intact nuclei. The interest of the UV cross-linking method is two-fold. First, it rules out the possibility that proteins could dissociate from physiological targets and artifactually reassociate with inappropriate targets during the experiment. Secondly, since the irradiation promotes a covalent attachment of proteins to DNA, the population of immunopurified fragments would comprise physiological

target sequences of various affinities for *Ubx* proteins. Homeoproteins indeed bind their targets with affinities presumably variable in strength depending on the bases flanking the ATTA core (Egger *et al.*, 1991), on interaction with other proteins (Smith and Johnson, 1992) and on chromatin structure. Differential affinity could be of important functional meaning taking into account that homeotic genes presumably control large sets of overlapping subordinate genes. The notion that homeoproteins compete for the control of common targets is suggested by several lines of data: similarities in DNA-binding specificity reported for several homeoproteins (Treisman *et al.*, 1992), transformations induced by the ectopic expression of homeotic genes (Mann and Hogness, 1990; Gonzalez-Reyes and Morata, 1991) and regulation of the expression of cloned targets by several homeotic genes (Gould *et al.*, 1990; this paper). It has been proposed (Gonzalez-Reyes *et al.*, 1990; Morata and Struhl, 1991) that differential affinity and competition for common downstream targets could account for the functional hierarchy existing between homeotic genes in determining the segmental identity. The isolation of such targets will provide clues for understanding the molecular basis of functional interactions between homeotic genes.

sca, a putative *in vivo* target of *Ubx*

We have provided three lines of evidence demonstrating that *sca* is an *in vivo* target of *Ubx*. First, the clone Cu8, unambiguously identified as the 480 bp *EcoRI* fragment that overlaps parts of last intron and exon of the *sca* gene, has been selected from the library of *Ubx* targets. Secondly, this fragment binds *Ubx* protein *in vitro*. Consistently, five motifs

that were required for homeodomain binding and presenting the ATTA core, were identified in the sequence of Cu8. Two of the three motifs situated in the 123 bp intronic sequence resemble the motif of *Ubx* binding found in the clone 48 isolated by Gould *et al.* (1990) and the consensus given by Desplan *et al.* (1988). This localization, as far as 10 kb from the promoter of the gene, suggests that the *cis-acting* sequences that bind *Ubx* proteins might be silencer elements; alternatively, the blockade of transcription might result from steric hinderance by the bound *Ubx* proteins. Thirdly, in stage 11 embryos, *sca* is transcribed in cell clusters in the three thoracic but not in the abdominal segments; this segmental pattern is complementary to the segmental pattern of *Ubx* proteins accumulation. In embryos deficient for *Ubx*, *sca* becomes ectopically expressed in the first abdominal segment, i.e. in the segment whose identity precisely requires *Ubx* functions. When taken together, these results suggest that *Ubx* proteins act as negative regulators of *sca* transcription in segment A1. As discussed by Andrew and Scott (1992), however, additional data will be required to definitely establish a status of direct *Ubx* target to *sca*.

We also found that *sca* is downregulated by *abdA* in segments A2–7 and by *AbdB* in segments A8–9, indicating that the three genes of BX-C are involved in its repression in the abdominal segments of the wild type embryo. *sca* is thus an example of a common target to several homeotic genes. Moreover, the expression of *Ubx* in segments A2–A7 of *abdA*-deficient embryos is sufficient to silence *sca* transcription. Similarly, the ectopic expression of *abdA* in segments A8 and A9 of *AbdB* deficient embryos is able to repress *sca*. The products of BX-C genes are, therefore, interchangeable for the repression of the common downstream gene, since *Ubx* can serve as a substitute for *abdA* and *abdA* for *AbdB*. While we have no molecular evidence for a direct interaction between the products of the *abdA* and *AbdB* genes, and regulatory sequences in *sca*, these observations could suggest either that the homeoproteins produced by BX-C genes bind competitively to the same site in the *sca* fragment identified by clone Cu8, or that each one binds a distinct motif presenting a ATTA core within this fragment, or elsewhere in *sca*.

The downstream targets of homeotic genes must specify some characteristics that are central in the processes of morphogenesis at a cellular level, such as the rate and orientation of mitosis, size and shape of cells, cell–cell communications and eventually expression of terminal differentiation products (Garcia-Bellido, 1975). The nature of the *sca* gene product, a secreted polypeptide with homology to fibrinogen and its likely function in cellular communications during neurogenesis (Baker *et al.*, 1990), are consistent with a role in morphogenesis. The segment to segment modulation of *sca* transcription, which is observed in stage 12 embryos and the downregulation achieved by BX-C genes in abdominal segments, suggest that *sca* could be involved in processes that specify morphogenetic traits in thoracic versus abdominal segments. A role in the morphogenesis of leg imaginal discs sensory organs could be inferred from the expression in small clusters of ventral cells in the thoracic segments. It is worth pointing out that homeotic transformations generated by *Ubx* affect the embryonic PNS: first, embryos deficient for *Ubx* have extra sensory organs (Keilin's organs) in the transformed segment (Hayes *et al.*, 1984); secondly, the ectopic expression of *Ubx* suppress the Keilin's organs formation

in the three thoracic segments (Mann and Hogness, 1990; Gonzalez-Reyes and Morata, 1991).

On the embryonic expression of *sca*

A number of mutations that result in the loss of function of genes involved in essential cellular or developmental processes and do not lead to observable phenotypes or lead to subtle or unexpected defects have been described as indicative for the existence of gene(s) encoding redundant function (e.g. Bieber *et al.*, 1989; Elkins *et al.*, 1990; Pereira *et al.*, 1992). During neurogenesis, *achaete* and *scute* are functionally redundant in certain aspects of CNS and PNS development (Ghysen and Dambly-Chaudière, 1988); the *hairy* protein and unknown gene products presumably have redundant functions during eye imaginal disc development (Brown *et al.*, 1991).

Animals that are homozygous deficient for the *sca* gene have no visible embryonic phenotype and are perfectly viable. Mlodzik *et al.* (1990) and Baker *et al.* (1990) pointed out as a salient feature the incomplete penetrance of the adult phenotypes, variable number and size of ommatidia and occasional duplication of sensory bristles. These authors proposed that there may exist a gene (or genes) with redundant function that can substitute for *sca*, so that no effect on embryogenesis and a relative incompleteness in adult phenotypes result if only *sca* function is lost. An attractive approach to test this proposal would be, as performed by Brown *et al.* (1991) with *hairy*, to analyse the phenotypic consequences of inducing *sca* ectopic expression during embryogenesis.

The tight pattern of *sca* transcription suggests that its product could be involved in cell signalling processes during embryonic neurogenesis. Prior to neuroblast segregation, the distribution of *sca* transcript displays a metameric modulation into 14 segmentally repeated clusters of neuroectodermal cells. Consistent with the proposal of Mlodzik *et al.* (1990), the *sca* protein might be required at that time in order to identify equivalent cells in competent neurogenic regions in each metamere. As segregation proceeds, *sca* RNA is abundant in progenitor cells of various neural lineages and at a reduced level only in neighbouring cells that adopt an epidermal fate. The *sca* product secreted by neural precursors could function as a molecular signal in a mechanism of lateral (Ghysen and Richelle, 1979; Simpson, 1990) or mutual (Heitzler and Simpson, 1991; Goriely *et al.*, 1991) inhibition, preventing the neighbouring cells to adopt a neural identity. Baker *et al.* (1990) have proposed that the *sca* protein could also be involved in lateral inhibition during the R8 photoreceptor cell differentiation in the developing eye. Such a mechanism, which results in the singling out of one cell from its neighbours, seems difficult to reconcile with the expression in cell clusters observed in the procephalon and in gnathal and thoracic segments (Figure 5). One possibility could be that the *sca* signal secreted by neural pioneers is differentially read depending on the fate of the receiving cell, either as a recognition signal by equivalent cells or as an inhibitory signal by cells committed to an epidermal fate.

Materials and methods

Flies

Oregon R was used as wild type standard. The alleles of homeotic genes used were *Ubx*^{9.22} (Kerridge and Morata, 1982), *abdA*^{MX1} (Sanchez-

Herrero *et al.*, 1985), *AbdB^{S4}* (Casanova *et al.*, 1986), *Df(3R)P9* and *Df(3R)109* (Lewis, 1978), *Antp^{W10}* (Wakimoto *et al.*, 1984).

In situ hybridization

Whole-mount hybridization with digoxigenin labelled probes was performed using the protocol of Tautz and Pfeifle (1989). The following probes were used to detect *sca* RNA: the 480 bp insert from Cu8 clone and a fragment of 992 nucleotides from the last exon of *sca* (this fragment was generated from genomic DNA by PCR amplification).

Library of genomic targets for Ubx proteins

Nuclei were isolated from 5–15 h old embryos in the presence of bivalent cations (5 mM Mg²⁺ and 2 mM Ca²⁺) and 0.5 M hexylene-glycol (Wray *et al.*, 1977).

Generation of protein–DNA adducts. UV cross-linking was essentially performed as described by Gilmour and Lis (1985). Nuclei were irradiated for 10 min at 4°C under a germicidal UV (254 nm) lamp delivering 4 × 10⁴ erg/cm²/s. After lysis in 4% lauroyl–sarcosine, CsCl was added to a final density of 1.3. The proteins floating at the surface were eliminated after centrifugation for 10 min at 20 000 g. The DNA was sheared by 10 passages through 1 mm inner diameter needle and loaded on the bottom of a CsCl step-density gradient (1.3, 1.5 and 1.7). The gradient was centrifuged for 17 h at 55 000 r.p.m. at 20°C in a 75 Ti rotor. The DNA fractions were pooled, dialysed twice against 2 l of dialysis buffer (10 mM Tris, 100 mM NaCl, 1 mM EDTA and 0.5 mM PMSF, pH 7.5), ethanol precipitated and resuspended in *Eco*RI digestion buffer. Digestion was performed in the presence of 0.8% NP 40 (10 units per µg of DNA, three times for 2 h each).

Immunoselection and cloning of Ubx protein adducts. *Eco*RI digested samples were preincubated for 1 h at 4°C with protein A–Sephrose beads. The preincubated samples were then cycled overnight at 4°C through a column of protein A–Sephrose beads coupled to monoclonal antibody FP3.38, in the presence of an excess of yeast tRNA to avoid non-specific binding of DNA. Sephrose beads were washed three times each with buffer I (10 mM Tris, 0.5 M LiCl, 0.1% SDS and 2% NP40, pH 7.4) and with buffer II (50 mM Tris and 50 mM NaCl, pH 8.0). The immunoprecipitate was eluted with 4% lauroyl–sarcosine, dialysed against 10 mM Tris buffer (pH 7.4) containing 1 mM EDTA, digested with proteinase K, extracted twice with phenol, concentrated with isobutanol and ethanol precipitated. The resulting DNA fragments were then cloned into λgt10.

In vitro binding to Ubx protein

The *Ubx* fusion protein construct 'pETUBX 11' was kindly provided by R.A.H. White (University of Cambridge). Heparin–agarose purified *Ubx* fusion protein extract (1 µg) was incubated for 15 min at 4°C in 50 µl of binding buffer (50 mM Tris, 200 mM NaCl, 0.1% Triton X100, 5 mM dithiothreitol, 10% glycerol, 1 mg/ml BSA and 0.5 mg/ml poly dI–dC, pH 7.5). ³²P-labelled insert DNA (10 ng per clone) and 1 µg of sonicated herring sperm DNA were added, and the incubation continued for 45 min. 100 µg of monoclonal antibody FP 3.38 conjugated to Trisacryl beads (IBF) were added to the binding reaction. After 90 min incubation, beads were washed three times with the binding buffer and the DNA was released by a stepwise increase of the salt concentration in the same buffer. The various fractions were analysed by electrophoresis on agarose gel (0.8%) and autoradiography.

DNA sequencing

Cu8 insert was cloned in the Bluescript KS⁺ vector (Stratagene), deletions generated with exonuclease III and the nucleotide sequence established by the dideoxy nucleotide technique of Sanger *et al.* (1977) using the sequenase II kit (United Biochemica Corp.).

Acknowledgements

We thank E. Alexandre, Y. Barthalay, A. Ghysen, S. Kerridge, R. Miassod, B. Oliver, M. Piovant and R. Rosset for critical reading of the manuscript and for helpful discussions. We acknowledge R.A.H. White for providing us with monoclonal antibody FP3.38 and *Ubx* fusion protein 'PETUBX 11'. This work was supported by the C.N.R.S. and a grant from A.R.C. to J. Pradel.

References

Akam, M. (1987) *Development*, **101**, 1–22.
Andrew, D.J. and Scott, M.P. (1992) *The New Biologist*, **4**, 5–15.

Baker, N.E., Mlodzik, M. and Rubin, G.M. (1990) *Science*, **250**, 1370–1377.
Beachy, P.A., Krasnow, M.A., Gavis, E.R. and Hogness, D.S. (1988) *Cell*, **55**, 1069–1081.
Bieber, A.J., Snow, P.M., Hortsch, M., Patel, N.H., Jacobs, J.R., Traquina, Z.R., Schilling, J. and Goodman, C.S. (1989) *Cell*, **59**, 447–460.
Bodmer, R., Caretto, R. and Jan, H.Y. (1989) *Neuron*, **3**, 21–32.
Brown, N.L., Sattler, C.A., Markey, D.R. and Carroll, S.B. (1991) *Development*, **113**, 1245–1256.
Busturia, A., Vernos, I., Macias, A., Casanova, J. and Morata, G. (1990) *EMBO J.*, **9**, 3551–3555.
Campos-Ortega, J.A. and Hartenstein, V. (1985) (eds) *The Embryonic Development of Drosophila melanogaster*. Springer Verlag, Berlin.
Carroll, S.B., DiNardo, S., O'Farrell, P.H., White, R.A.H. and Scott, M.P. (1988) *Genes Dev.*, **2**, 350–360.
Casanova, J., Sanchez-Herrero, E. and Morata, G. (1986) *Cell*, **42**, 663–669.
Celniker, S.E., Keelan, D.J. and Lewis, E.B. (1989) *Genes Dev.*, **3**, 1424–1436.
Champlin, D.T., Frasch, M., Saumweber, H. and Lis, J.T. (1991) *Genes Dev.*, **5**, 1611–1621.
Chatterjee, P.K., Bruner, M., Flint, S.J. and Harter, M.L. (1988) *EMBO J.*, **7**, 835–841.
Desplan, C., Theis, J. and O'Farrell, P.H. (1988) *Cell*, **55**, 1081–1090.
Dessain, S., Gross, C.T., Kuziora, M.A. and McGinnis, W. (1992) *EMBO J.*, **11**, 991–1002.
Ekker, S.C., Young, K.E., Von Kessler, D.P. and Beachy, P.A. (1991) *EMBO J.*, **10**, 1179–1186.
Elkins, T., Ganetsky, B. and Wu, C.F. (1990) *Proc. Natl. Acad. Sci. USA*, **83**, 8415–8419.
Fitzpatrick, V.D., Percival-Smith, A., Ingles, C.J. and Krause, H.M. (1992) *Nature*, **356**, 610–612.
Garcia-Bellido, A. (1975) *Ciba Found. Symp.*, **29**, 161–178.
Garzino, V., Moretti, C. and Pradel, J. (1987) *J. Biol. Cell.*, **61**, 5–14.
Garzino, V., Berenger, H. and Pradel, J. (1989) *Development*, **106**, 17–27.
Gavis, E.R. and Hogness, D.S. (1991) *Development*, **112**, 1077–1093.
Gehring, W.J. (1987) *Science*, **236**, 1245–1252.
Ghysen, A. and Richelle, J. (1979) *Dev. Biol.*, **70**, 438–452.
Ghysen, A. and Dambly-Chaudière, C. (1988) *Genes Dev.*, **2**, 495–501.
Ghysen, A. and Dambly-Chaudière, C. (1989) *Trends Genet.*, **5**, 251–255.
Ghysen, A. and O'Kane, C. (1989) *Development*, **105**, 35–52.
Gilmour, D.S. and Lis, J.T. (1985) *Mol. Cell. Biol.*, **5**, 2009–2018.
Gilmour, D.S., Pflugfelder, G., Wang, J.C. and Lis, J.T. (1986) *Cell*, **44**, 401–407.
Gonzalez-Reyes, A., Urquia, N., Gehring, W.J., Struhl, G. and Morata, G. (1990) *Nature*, **344**, 78–80.
Gonzalez-Reyes, A. and Morata, G. (1991) *Development*, **113**, 1459–1471.
Goriely, A., Dumont, N., Dambly-Chaudière, C. and Ghysen, A. (1991) *Development*, **113**, 1395–1404.
Gould, A.P., Brookman, J.J., Strutt, D.I. and White, R.A.H. (1990) *Nature*, **348**, 308–312.
Hafen, E., Levine, M. and Gehring, W.J. (1984) *Nature*, **307**, 287–289.
Harding, K., Wedeen, C., McGinnis, W. and Levine, M. (1985) *Science*, **229**, 1236–1242.
Hartenstein, V. and Campos-Ortega, J.A. (1984) *Wilhelm Roux's Arch. Dev. Biol.*, **193**, 308–325.
Hayes, P.H., Sato, T. and Denell, R.E. (1984) *Proc. Natl. Acad. Sci. USA*, **81**, 545–549.
Heitzler, P. and Simpson, P. (1991) *Cell*, **64**, 1083–1092.
Hoey, T. and Levine, M. (1988) *Nature*, **332**, 858–861.
Immerglück, K., Lawrence, P.A. and Bienz, M. (1990) *Cell*, **62**, 261–268.
Ingham, P. (1988) *Nature*, **335**, 25–34.
Ingham, P.W. and Martinez-Arias, A. (1992) *Cell*, **68**, 221–236.
Johnson, F.B. and Krasnow, M.A. (1990) *Genes Dev.*, **4**, 1044–1052.
Kerridge, S. and Morata, G. (1982) *J. Embryol. Exp. Morph.*, **68**, 211–234.
Klambt, C., Jacobs, J.R. and Goodman, C.S. (1991) *Cell*, **64**, 801–816.
Lewis, E.B. (1978) *Nature*, **276**, 565–570.
Mann, R.S. and Hogness, D.S. (1990) *Cell*, **60**, 597–610.
McGinnis, W. and Krumlauf, R. (1992) *Cell*, **68**, 283–302.
Mlodzik, M., Baker, N.E. and Rubin, G.M. (1990) *Genes Dev.*, **4**, 1848–1861.
Morata, G. and Struhl, G. (1990) *Nature*, **348**, 587–588.
Nüsslein-Volhard, C.H., Frohnhofer, H.G. and Lehmann, R. (1987) *Science*, **238**, 1675–1681.
Paine, P.L., Austerberry, C.F., Desjarlais, L.J. and Horowitz, S.B. (1983) *J. Cell Biol.*, **197**, 1240–1243.
Patel, N.H., Snow, P.M. and Goodman, C.S. (1987) *Cell*, **48**, 975–988.
Peifer, M. and Wieschaus, E. (1990) *Genes Dev.*, **4**, 1209–1223.
Pereira, A., Doshen, J., Tanaka, E. and Goldstein, L.S.B. (1992) *J. Cell Biol.*

- 116, 377–383.
- Reuter, R., Panganiban, G.E.F., Hoffmann, F.M. and Scott, M.P. (1990) *Development*, **110**, 1031–1040.
- Sanchez-Herrero, E., Vernos, I., Marco, R. and Morata, G. (1985) *Nature*, **313**, 108–113.
- Sanger, F., Nicklen, S. and Coulson, A.R. (1977) *Proc. Natl. Acad. Sci. USA*, **74**, 5463–5467.
- Scott, M.P., Tamkun, J.W. and Hartzell, G.W. (1989) *Biochim. Biophys. Acta*, **989**, 25–48.
- Simpson, P. (1990) *Development*, **109**, 509–519.
- Smith, D.L. and Johnson, A.D. (1992) *Cell*, **68**, 133–142.
- St Johnston, D. and Nüsslein-Volhard, C. (1992) *Cell*, **68**, 201–220.
- Struhl, G. and White, R.A.H. (1985) *Cell*, **43**, 507–519.
- Tautz, D. and Pfeifle, C. (1989) *Chromosoma*, **98**, 81–85.
- Thali, M., Müller, M.M., De Lorenzi, M., Matthias, P. and Bienz, M. (1988) *Nature*, **336**, 598–601.
- Thomas, J.B., Bastiani, N.J., Bate, M. and Goodman, C.N. (1984) *Nature*, **310**, 203–207.
- Treisman, J., Harris, E., Wilson, D. and Desplan, C. (1992) *BioEssays*, **14**, 145–150.
- Wagner-Bernholz, J.T., Wilson, C., Gibson, G., Schuh, R. and Gehring, W.J. (1991) *Genes Dev.*, **5**, 2467–2580.
- Wakimoto, B.T., Turner, F.R. and Kaufman, T.C. (1984) *Dev. Biol.*, **102**, 147–172.
- White, R.A.H. and Wilcox, M. (1984) *Cell*, **39**, 163–171.
- Wray, W.W., Conn, M. and Wray, J.P. (1977) *Methods Cell Biol.*, **16**, 69–86.

Received on February 27, 1992; revised on May 18, 1992